Genetics of heavy metal hyperaccumulation in Thlaspi caerulescens

Deniau, A.X.

2010

document version
Publisher's PDF, also known as Version of record

Genetics of heavy metal hyperaccumulation

in

Thlaspi caerulescens

Antoine Xavier Deniau

The research presented in this thesis was carried out at the Vrije Universiteit-Amsterdam and partly at the Wageningen Universiteit in the Netherlands.

This research was financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek-Genomics (NWO) Grant (050-10-166).

Front cover: DNA ligase (shown in color, encircling the DNA double helix).

Back cover: also DNA ligase (shown in color, encircling the DNA double helix).

VRIJE UNIVERSITEIT

Genetics of heavy metal hyperaccumulation in *Thlaspi caerulescens*

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Aard- en Levenswetenschappen
op maandag 13 september 2010 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Antoine Xavier Deniau

geboren te Nantes, Frankrijk

promotor: prof.dr. R.E. Koes
copromotor: dr. H. Schat

“Un fruit est une fleur qui a connu l’amour”
anonyme, INRA

Table of contents

Chapter 1

General Introduction

Research on hyperaccumulators

A relatively small number of plant species, generally referred to as metal hyperaccumulators, can accumulate very high levels of essential and non-essential metals (Baker & Brooks 1989). Metal-hyperaccumulating plants (Brooks *et al.,* 1977) have been defined as plants that accumulate more than 1000 μg g^{-1} nickel, 10 000 μg g^{-1} zinc, or 100 μg g^{-1} cadmium in their foliage when growing in nature (Reeves, 1988; Reeves, 1992; Baker & Brooks, 1989; Baker *et al.,* 2000). The perspective of phytoremediation has greatly increased the interest for metal-hyperaccumulator plants due to their potential use in the clean-up of metal-polluted soils (Pollard *et al.*, 2002; Baker *et al.,* 1991). Much progress has been made over the last decade in the understanding of the physiology and molecular mechanisms of metal hyperaccumulation (see reviews by Pollard *et al.,* 2002; Macnair, 2003; McGrath & Zhao, 2003; Verbruggen *et al*., 2009). However, a robust evolutionary explanation of the hyperaccumulation trait is still lacking. Boyd & Martens (1992) and Boyd (1998) discussed several hypotheses concerning the possible selective advantage of hyperaccumulation. These hypotheses include: inadvertent uptake, metal tolerance/disposal, drought resistance, interference with other plants (elemental allelopathy), and defence against pathogens or herbivores. The "metal defence" hypotheses is among the most attractive hypothesis for the "reason to be" of metal hyperaccumulation and appears to be supported by some experimental data (Boyd *et al.,* 2002 & 2007; Hanson *et al.,* 2003; Huitson & Macnair, 2003; Jiang *et al.,* 2005; Noret *et al.,* 2005). Metals can protect plants from herbivory and pathogen infection in different ways, acting either as

an antifeedant or as a plant systematic pesticide (Poschenrieder *et al.,* 2006). However, if inadvertent uptake is indeed the ultimate reason for heavy metal hyperaccumulation, then still enhanced defence against herbivory may be a consequence. This benefit could conceivably reinforce the selection for heavy metal hyperaccumulation.

Among the Brassicaceae, *Thlaspi caerulescens* J. & C. Presl and *Arabidopsis halleri* are considered to be attractive genetic model species because of their physiological, morphological and genetic characteristics (Pollard *et al*., 2002; Assunção *et al.,* 2003a; Macnair, 2003; Peer *et al.,* 2003; Peer *et al.,* 2006). An additional advantage is their close relationship with the universal plant genetic model, *Arabidopsis thaliana* L. Heynh. *T. caerulescens* shares on average 88.5% DNA identity in coding regions (Rigola *et al.,* 2006) and 87% DNA identity in the intergenic transcribed spacer regions (Peer *et al.,* 2003) with *A. thaliana*. *A. halleri* and *A. thaliana* share on average 94% nucleotide sequence identity within coding regions (Becher *et al.,* 2004). Most research on hyperaccumulators has been focussed on the physiological mechanisms of metal uptake, transport and sequestration (McIntyre, 2003; Long *et al.* 2002). Despite these efforts relatively little is known regarding the genetic basis and molecular mechanisms of hyperaccumulation (Macnair *et al.,* 2000). Thus, diverse "–omics" approaches have been used to tackle the problem, such as transcriptomics, proteomics and genomics, but also the more classical genetic approaches of segregation and co-segregation analysis of intra- and inter-specific crosses, and ultimately quantitative trait loci (QTL) mapping (Assunção *et al*., 2006; Courbot *et al.*, 2007; Willems *et al.,* 2007). The QTL analysis of the *A. halleri* x *A.*

lyrata (hyperaccumulator x non-hyperaccumulator) BC1 yielded candidates, HMA4 for Cd and Zn tolerance, and MTP1a and MTP1b for Zn tolerance, respectively. In *T. caerulescens*, which can not be crossed to non-hyperaccumulating congeneric species, these research strategies benefit from the pronounced intra-specific variation in accumulation and tolerance characteristics among local populations (Assunção *et al*., 2003b & 2003c). The diverse "-omics" approaches will be discussed with a special emphasis on *T. caerulescens*. The aim of this introduction is to sketch the background and the framework in which this thesis fits and what were the starting points.

Transcriptomics

Genome-wide expression analysis is a powerful tool to obtain clues about the genes that are involved in adaptive responses to environmental conditions. Unfortunately, "whole-genome" micro-arrays have not been developed for metal hyperaccumulating species yet, because of the high costs associated with their design and production (van de Mortel & Aarts, 2006a). For *T. caerulescens*, spotted cDNA arrays have been developed (Plessl *et al.,* 2005). The number of genes represented on such arrays is limited. More information has been obtained through hybridisation to heterologous, but genome-wide, microarrays, as has been performed with *A. halleri* (Weber *et al.,* 2004; Becher *et al*., 2004) and *T. caerulescens* (Hammond *et al.,* 2006; van de Mortel *et al.,* 2006b).

By means of comparative transcript profiling of metal hyperaccumulator and non-accumulator shoots and roots, Weber *et al.* (2004) identified a

distinct expression and organ-specific regulation of different members of the *ZIP* and *NAS* metal homeostasis gene families in the zinc hyperaccumulator *A. halleri*, as compared to *A. thaliana.* This reflects altered cellular functions common to roots and shoots in hyperaccumulators. These functions include an elevated cellular zinc uptake rate, with a predominant role of *ZIP9* in roots and *ZIP6* in shoots, and an enhanced low-molecular weight chelator biosynthesis to achieve sufficient cytoplasmic buffering and intercellular metal mobility, with major roles for *NAS2* in roots and possibly *NAS3* in shoots. The striking differences between root and shoot transcript profiles may reflect the different functions of the root and the shoot in metal hyperaccumulation and tolerance (Guimaraes *et al.,* 2009). In the roots of *A. halleri*, *NAS2* and the natural resistance-associated macrophage protein family member (*NRAMP3*), which encodes a protein involved in iron, manganese, cadmium, and possibly, zinc transport, are highly expressed and likely to have a role in sustaining root-to-shoot mobility of zinc (Thomine *et al.,* 2000; Thomine *et al.,* 2003; Filatov *et al*., 2006). Becher *et al.* (2004) suggest that *MTP1* and *HMA3*, which are predominant in the shoot transcript profiles, are important for metal sequestration and detoxification, thereby generating a metal sink in the shoot that could be one of the driving forces in metal hyperaccumulation. However, as indicated by reciprocal grafting experiments, hyperaccumulation may not be dependent on the sink strength of the shoot (Guimaraes et al., 2009). Recently, using a combination of genome-wide cross species microarray analysis (ATH1 GeneChip® array) and real-time RT-PCR, Talke *et al.* (2006) identified a set of candidate genes for zinc

hyperaccumulation, zinc and cadmium hypertolerance, and the adjustment of micronutrient homeostasis in *A. halleri*. Eighteen putative novel metal homeostasis genes were found to be more expressed in *A. halleri* than in *A. thaliana*, and 11 previously identified candidate genes were confirmed. The encoded proteins include HMA4, known to contribute to root-shoot transport of zinc in *A. thaliana* (Hussain *et al.,* 2004; Verret *et al.,* 2004; Mills *et al.,* 2005). Expression of either *AtHMA4* or *AhHMA4* confers cellular zinc and cadmium tolerance in yeast (*Saccharomyces cerevisiae*) (Talke *et al*. 2006). More recently, *AhHMA4* was shown to be much higher expressed than *AtHMA4*, due to a triplication of the harbouring genomic region, as well as altered *cis*-regulation. Expression of *AhHMA4* in *A. thaliana* enhanced Zn accumulation in the shoot and increased Zn tolerance. On the other hand, silencing of *AhHMA4* reduced Zn and Cd tolerance and decreased the translocation of these metals to the shoot in *A. halleri* to the level observed in wild-type *A. thaliana* (Hanikenne *et al.*; 2008). Among the genes identified by Talke *et al.* (2006) were *IRT3* and *ZIP10*, which have been proposed to contribute to cytoplasmic zinc influx, and *FRD3* which is required for iron transport in *A. thaliana* (Green & Rogers 2004). In *A. halleri*, the presence of more than a single genomic copy is a hallmark of several highly expressed candidate genes with possible roles in metal hyperaccumulation, i.e. *HMA4* and *MTP1* (Talke *et al.* 2006). Zinc hyperaccumulation in *A. halleri* involves enhanced translocation of zinc from roots into shoots. The transcriptional regulation of marker genes suggests that in the steady state, *A. halleri* roots, but not the shoots, act as physiologically zinc deficient under conditions of moderate zinc supply

(Talke *et al*. 2006). Zn deficiency in the root could be created by the high rate of the metal translocation to the shoot mediated by HMA4 (Hanikenne *et al.,* 2008). However, highly metal hyperaccumulating *T. caerulescens* accessions can accumulate metals to higher levels in the roots than non-accumulators do, even at relatively low heavy metal exposure levels, which argues against this hypothesis (Richau *et al.,* 2009).

The transcriptome of *T. caerulescens* has been profiled using a custom cDNA spotted microarray representing 1900 expressed sequence tags (ESTs) and comparing accessions from La Calamine (LC) and Lellingen (LE) of *T. caerulescens* grown under different zinc exposure conditions (Plessl *et al.,* 2005). Although differences were observed between the two accessions and between metal exposure conditions, these differences were not very large, despite the differences in zinc uptake and tolerance of the tested accessions. Most remarkable was the difference in the expression of genes involved in synthesis of nicotianamine precursors, which were much higher expressed in LC compared to LE, at elevated zinc concentrations (100 and 1000 μM). Hammond *et al.* (2006) provided a more comprehensive comparison of the transcriptional profiles of *T. caerulescens* and *T. arvense* shoots. This species couple has often been used in comparative physiological and molecular studies on metal hyperaccumulators and non-hyperaccumulators (Lasat *et al.,* 1996; Pence *et al.,* 2000; Assunção *et al.,* 2001; Pineros & Kochian, 2003). In this study, a newly described gDNA-based probe-selection strategy was used to profile and compare the transcriptomes of *T. caerulescens* and *T. arvense*. This strategy was shown to be robust and

valid, based on *in silico* alignments of ATH1 GeneChip® array probes with *T. caerulescens* and *T. arvense* sequences, and quantitative real-time PCR. In total, 4947 transcripts (representing homologues, possibly orthologues, of genes in *A. thaliana*) were identified as differentially (> 2-fold or < 0.5-fold) expressed in the shoots of *T. caerulescens* compared with *T. arvense*. In the shoots the abundance of 3349 transcripts was higher in *T. caerulescens* than in *T. arvense* and the abundance of 1598 transcripts was lower. Many transcripts encoding proteins with putative roles in cellular zinc homeostasis were differentially expressed in the shoots of *T. caerulescens* compared to *T. arvense*, including several genes previously shown to be involved in zinc transport and zinc compartmentalization. There was differential expression of five genes homologous to the *A. thaliana* ZIP transporter family, including higher expression of *AtIRT3*, *AtZIP6* and *AtZIP7*, and lower expression of *AtZIP3* and *AtZIP10* in shoots of *T. caerulescens* compared with *T. arvense*. Transcripts with homology to three CDF transporters (At2g39450, *AtMTP11*; At2g04620, *AtMTP12*; At3g12100, *AtMTP5*; Delhaize *et al.,* 2003) were expressed at higher levels in *T. caerulescens* than in *T. arvense*. The transcript abundance of *TcMTP1* has previously been shown to be higher in the roots and shoots of three *T. caerulescens* ecotypes than in the non-hyperaccumulator *T. arvense* (Assunção *et al.,* 2001). Remarkably, *TcMTP1* (At2g46800) transcript abundance was lower in *T. caerulescens* than in *T. arvense* in the study of Hammond *et al.* (2006). Also, four genes of the P-type ATPase family had higher expression in shoots of *T. caerulescens* compared with *T. Arvense*, i.e. the homologues of the P1B-type ATPases *AtHMA3* and

AtHMA4, and the Ca^{2+}-transporting ATPases *AtACA12* and *AtACA13*. Of the genes identified in the studies of Becher *et al.* (2004) and Weber *et al.* (2004) utilizing the *A. thaliana* AG GeneChip® array, only 16 genes were common to the genes identified as significantly differentially expressed between *T. caerulescens* and *T. arvense* in the study of Hammond *et al.* (2006). These included transcripts homologous to *AtCAX2*, *AtHMA3*, *AtZIP6*, *AtZAT/MTP1*, and a cytochrome P450 transcript.

More recently, using a 60-mer oligo microarray containing 40 000 probes that were designed to represent full-genome coverage of *A. thaliana* (Arabidopsis 3 oligo microarray; Agilent Technologies Inc.), van de Mortel *et al.* (2006b) have examined in detail the transcription profiles of roots of *A. thaliana* and *T. caerulescens* plants grown under three different conditions: deficiency, sufficiency, and excess of zinc. A total of 608 zinc-responsive genes with at least a 3-fold difference in expression level in response to changes in zinc supply were detected in *A. thaliana* and 352 in *T. caerulescens*. Only 14% of the genes that were zinc responsive in *T. caerulescens*, were also zinc responsive in *A. thaliana*. When comparing *A. thaliana* with *T. caerulescens* at each of the zinc exposure levels, more than 2,200 genes were significantly differentially expressed ($\geq$ 5-fold at a false discovery rate < 0.05). While a large fraction of these genes are of yet unknown function, many genes with a different expression between *A. thaliana* and *T. caerulescens* appear to function in metal homeostasis. Van de Mortel *et al.* confirmed that a number of genes previously known (*ZIP2*, *4*, *5* and *9*, *NAS2* and *HMA2*) or suggested (*ZIP1*, *3*, and *10*, *IRT3*, *MTP2* and *NAS4*) to be

involved in zinc homeostasis were induced under low zinc supply. Their results suggest that the ZIP gene family plays an important role in zinc uptake in the root, with 10 members differentially expressed in this organ (*ZIP1*, *2*, *3*, *4*, *5*, *9*, *10*, *11*, and *12*, and *IRT3*). The authors propose that these transporters exert a similar function in different parts of the root, or are located at different membranes. Rather than *MTP1* (previously known as *ZAT*), which is constitutively expressed in Arabidopsis (van der Zaal *et al.,* 1999; Kobae *et al*., 2004), the induction of *MTP2* by zinc deficiency suggests a specific role of this transporter in counteracting the effect of zinc deficiency. Only two of the Arabidopsis *NAS* genes (*NAS2* and *NAS4*) are more expressed in roots under zinc deficiency compared to sufficiency. The presence of four, apparently paralogous NAS genes with different, partly overlapping, gene expression profiles suggests complementarity, although there may be a degree of redundancy. The strong expression of *NAS2* in *A. halleri* compared to *A. thaliana* (Weber *et al.*, 2004; Talke *et al.,* 2006) was not found in *T. caerulescens*, but instead *TcNAS1* and *TcNAS3*, and to a lesser extent *TcNAS4*, were higher expressed, particularly in shoots of *T. caerulescens* (van de Mortel *et al*., 2006b). The differential expression profiles of three NAS genes in *T. caerulescens*, compared to *A. thaliana*, suggest a major function of these nicotianamine synthase genes in metal adaptation of *T. caerulescens*. In addition to the higher expression of the NAS genes, van de Mortel *et al.* (2006b) found an unexpectedly high expression of *FRO5*, *FRO4* and *FRD3* under zinc deficiency. *FRO4* and *FRO5* resemble the ferric chelate reductase gene *FRO2* (Robinson *et al.,* 1999) but, in contrast to *FRO2*, their expression is not induced in *A.*

thaliana roots upon iron deficiency (Mukherjee *et al.,* 2005; Wu *et al,.* 2005). Although *FRD3* has been mainly implicated in iron homeostasis (Green and Rogers, 2004), the gene has also been identified in the study of Talke *et al.* (2006). This suggests that iron homeostatic genes may have a much broader role in general metal homeostasis than previously thought. When considering the 16 highest expressed genes of a set of genes that is at least five-fold higher expressed in *T. caerulescens* compared to *A. thaliana* at zinc sufficiency, already six metal homeostasis genes are among them, four of which are known zinc transporters: *HMA4*, originally isolated by Bernard *et al.* (2004) and characterised by Papoyan & Kochian (2004) as a zinc-transporting P-type ATPase possibly involved in the loading of zinc into the xylem; *MTP1*, described as *ZTP1* by Assunção *et al.* (2001); *ZIP4* and *IRT3*, described as *TcZNT1* and *TcZNT2* by Assunção *et al.* (2001). Other zinc transporter genes that are more expressed in *T. caerulescens* are *HMA3*, a P-type ATPase similar to *HMA4*; *MTP8*, a member of the cation diffusion facilitator (CDF) family; and *NRAMP3*. *HMA3* has also been found to be differentially expressed in the studies of Becher *et al.* (2004), Weber *et al.* (2004) and Hammond *et al.* (2006), suggesting an important role of this gene in metal homeostasis. Also differentially expressed in *A. halleri*, *NRAMP3* is likely to have a role in sustaining root-to-shoot mobility of zinc (Thomine *et al.,* 2003, Filatov *et al.,* 2006).

The comparative transcriptomic analysis of the hyperaccumulator *T. caerulescens* and the nonaccumulators *A. thaliana* or *T. arvense* emphasizes the role of previously implicated zinc homeostasis genes but also suggests a similar role for many more, as yet uncharacterised genes.

While some of these genes were also differentially expressed in *A. halleri* - *A. thaliana* comparisons, many more were not, suggesting that there is some overlap in the mechanisms of metal hyperaccumulation but that there are also many differences between different metal hyperaccumulator species.

Although a *T. caerulescens* specific cDNA array has been made (Plessl *et al.,* 2005), the number of cDNA sequences spotted is still far from that needed to make a "whole-genome" cDNA array (Rigola *et al.,* 2006). Non-targeted approaches, such as differential display (DD) (Liang & Pardee, 1992), are viable alternatives for intra-specific comparisons, at least. As demonstrated by Mandaokar *et al.* (2003), additional differentially expressed genes were found from *A. thaliana* with DD compared to microarray. DD has been applied for the isolation of genes involved in many processes in Arabidopsis, such as the response to cadmium (Suzuki *et al.,* 2001). Hassinen *et al.* (2006) compared shoot transcript patterns of two *T. caerulescens* accessions using DD. Two zinc-responsive *MT* genes in *T. caerulescens*, *TcMT2a* and *TcMT3*, apparently involved in intracellular metal binding, were identified. The authors suggested that enhanced expression of *TcMT2a* reflects an increased requirement for zinc binding. The *TcMT3* expression levels appeared to reflect the shoot zinc levels and may thus have a function in metal homeostasis, particularly copper homeostasis, at high zinc accumulation levels (Roosens *et al*., 2004, 2005). Two more genes with possible roles in metal sequestration were isolated by Hassinen *et al.* (2006), one encoding an MRP transporter, the other a pectine methylesterase (PME). The *MRP* gene was highly homologous to the

AtMRP10 and *ATMRP4* genes. The *AtMRPs* belong to a family of membrane-associated gluthatione-conjugate transporters, which form a superfamily of ABC transporters (Bovet *et al.,* 2003). ABC transporters are involved in the vacuolar sequestration of cadmium (Ortiz *et al.,* 1992; Wemmie *et al.,* 1994). In a yeast complementation study (Hassinen et al., 2006), the cadmium content was not altered, suggesting that *TcMRP* was not able to sequester cadmium in the vacuole. The authors proposed that this might be due to degradation or to the inappropriate targeting of the protein in the yeast. Recently, Kim *et al*. (2006) suggested that *MRP3* could be responsible for vacuolar Cd sequestration. In *T. caerulescens,* zinc is mainly stored in the vacuole (Kupper *et al.,* 1999), but the apoplast is also a major storage compartment (Frey *et al.,* 2000). PMEs modify the properties of cell walls by demethylation of pectin residues, thus creating free carboxylic groups for interaction with divalent cations. A higher expression of PME in the zinc-accumulator accession may thus enhance zinc binding in the apoplast. Another gene identified by DD was the *Ole e 1*-like gene (Ledesma *et al.,* 2006) under zinc deficiency in root. Hassinen *et al.* (2006) suggest that the *TcOle e 1* is not directly involved in cadmium or intercellular binding. The protein may contribute in a complex way to the maintenance of metal homeostasis or to the energy or metabolite balance during metal excess. In conclusion, even though a small selection of genes was identified using DD, (novel) genes with possible implications in metal adaptation were isolated.

Proteomics

The major advantage of proteomics over transcriptomics is that it focuses on the translated portion of the genome. The importance of post-transcriptional regulation has been underlined by several studies in yeast showing only a weak or moderate correlation between mRNA and protein levels, except for very abundant proteins (Gygi *et al.,* 1999; Ideker *et al.,* 2001). Proteomic profiling has been used to study the effects of several biotic and abiotic stress factors on plants. Only a few reports on the effects of metal ions on non-hyperaccumulator species proteomes are available to date (Thiellement *et al.,* 1999; Thiellement *et al.,* 2002). A study on the rice proteome indicated that some metals disrupt the photosynthetic machinery (Hajduch *et al.,* 2001). Also, Repetto *et al.* (2003) have shown that cadmium-induced changes in the pea root proteome were modulated by mycorrhizal symbiosis. In both studies only a limited number of proteins were clearly affected by exposure to heavy metals. Besides differences in protein spot intensity, reflecting differential expression, there may be allelic diversity, leading to different protein isoforms, with altered electrophoretic mobility. The fraction affected by heavy metal exposure is relatively low, compared to the fraction of transcripts affected, due to post-regulation and the lower sensitivity of protein profiling.

To my knowledge, the study of Ingle *et al.* (2005) is the first attempt to use proteomic profiling to explore the molecular mechanisms related to heavy metal accumulation in a hyperaccumulator species. These authors tried to identify proteins that play a role in nickel accumulation and tolerance in the nickel hyperaccumulator *Alyssum lesbiacum* by

analysing changes in protein abundance occurring in response to short- and long-term exposure to nickel. Short-term exposure led to a change in the abundance of several proteins associated with the metabolism of sulphur containing amino acids. Together these changes indicated a re-allocation of sulphur towards increased production of cystein. The increase in proteins such as oxidoreductases, a mannitol-6-phosphate reductase and a glutathione S-transferase may reflect an antioxidant system to prevent membrane damage, especially when the nickel concentration is close to the maximum that the plants can tolerate without growth inhibition. In the light of the small number of proteins seen to alter in abundance after a long-term exposure to nickel, Ingle *et al.* have made an attempt to compare the proteome of *A. lesbiacum* with that of *A. montanum* (a related non-accumulator) in the absence of added nickel. Unfortunately, the protein patterns of the two species were insufficiently similar to allow any analysis. This clearly suggests that the proteomic approach may only be successful in case of intra-specific comparisons.

Recently, Tuomainen *et al.* (2006) successfully compared protein patterns of three metal hyperaccumulator accessions of *T. caerulescens* with distinct characteristics of metal uptake and transport when exposed to different metal exposures. The strongest differences were seen when comparing the different accessions, rather than the different metal exposures. The 48 tentatively differential spots represented proteins with core metabolic functions (e.g. photosynthesis, nitrogen assimilation, carbohydrate metabolism) as well as putative signalling and regulatory proteins. Tuomainen *et al.* suggested that phenotyped crosses of the

Thlaspi populations might help in finding a possible connection to metal accumulation or tolerance. Nevertheless, these proteome data provide an important contribution to the systems biology approach to the metal hyperaccumulation phenomenon.

Evolutionary history of heavy metal hyperaccumulation and tolerance in *Thlaspi caerulescens*

Contrasting with the strong genetic differences in adaptive and life history traits, Jiménez-Ambriz *et al.* (2006) did not find any significant differentiation between *T. caerulescens* ecotypes for microsatellite markers. Analysis of variation for microsatellite loci, which are reputed to be highly polymorphic, thus confirms the findings of allozyme studies in *T. caerulescens* (Dubois *et al.,* 2003; Koch *et al.,* 1998), that showed a clear differentiation between accessions in Belgium and Luxembourg but not in southern France, in line with the lower degree of geographical isolation between the two accessions in this region. Despite the small geographic distances between populations in this study, they found significant differentiation for microsatellite markers among populations within edaphic origins (metallicolous or non-metallicolous). The high F_{SC} value (almost 20%), together with the absence of significant isolation by distance, suggests that gene flow, through either pollen or seeds, is very limited among different populations, confirming the conclusions of Dubois *et al.* (2003) and Koch *et al.* (1998). Gene flow might, therefore, provide only weak opposition to the evolution of local adaptations in the heterogeneous network of populations considered in

Jiménez-Ambriz *et al.* (2006). Selection at sites contaminated with heavy metals has been reported to be strong and to produce a rapid response (Reznick & Ghalambor, 2001). Local adaptation can evolve in a heterogeneous environment in the presence of genotype–environment interaction for fitness, strong selection and moderate gene flow between differentiating populations (Kawecki & Ebert, 2004). All of these conditions are met in the study of Jiménez-Ambriz *et al.* (2006). Metallicolous and nonmetallicolous plants showed very different phenotypic responses to experimental zinc exposure in a garden experiment. Gene flow between populations was low and their results suggest that strong divergent selection has shaped the patterns of ecotypic variation in heavy metal hyperaccumulation and tolerance.

Genetics

Many studies on metallophyte species have dealt with the genetics of heavy metal tolerance. So-called "pseudometallophytes" or "facultative metallophytes" occur both on metalliferous and normal soils. In these species high-level metal tolerance is usually metal-specific and confined to metallicolous populations. A number of investigators have made intra-specific crosses between metallicolous and non-metallicolous plants in order to study the segregation of metal tolerance (Urquhart, 1971; Macnair, 1983; Schat & Ten Bookum, 1992; Macnair, 1993). In contrast with heavy metal tolerance, data on the genetics of hyperaccumulation are scarce. The two phenomena – accumulation and tolerance – are largely independent of each other: tolerance is a phenomenon in which the plant interacts with the external metal concentration, with the internal

concentration regulated by homeostatic mechanisms with enhanced capacities in tolerant plants, compared to non-tolerants. Accumulation is a phenomenon primarily concerned with the distribution of metal over roots and shoots, though an increased uptake at the root level may also be found (Lasat *et al.,* 1996).

Despite the relative lack of tolerance to zinc in nonmetallicolous populations, all the populations of *T. caerulescens* hyperaccumulate this metal and thus are able to detoxify it to some extent (Assunção *et al.,* 2001); indeed, under controlled conditions, nonmetallicolous populations accumulate zinc to higher foliar concentrations than metallicolous populations, suggesting that adaptation to metalliferous soil involves a partial loss of the Zn accumulation capacity, probably due to the prevailing high concentrations of plant-available Zn in the soil (Escarré *et al.,* 2000; Dechamps *et al.,* 2007). Pollard & Baker (1996), who studied two metallicolous populations of *T. caerulescens* in Great Britain, showed a continuous heritable variation in zinc hyperaccumulation ability within both populations. The observed variation was suggested to be polygenically controlled (Pollard, 2000). The number of genes controlling the clear-cut difference in hyperaccumulation ability between metallicolous and nonmetallicolous populations, found by Meerts & Van Isacker (1997) and Escarré *et al.* (2000) is unknown. As far as the zinc hyperaccumulation ability is concerned, the results of Frérot *et al.* (2003) imply that the difference between metallicolous and nonmetallicolous populations is heritable and that the restricted degree of zinc hyperaccumulation in metallicolous parents is largely dominant over the higher degree of zinc

hyperaccumulation in nonmetallicolous parents. Frérot *et al.* (2003) claimed a monogenic system with two alleles could be responsible for the difference in zinc hyperaccumulation ability between metallicolous and nonmetallicolous populations, but their evidence is not conclusive.

In contrast to zinc hyperaccumulation, cadmium hyperaccumulation does not seem to be constitutive in *T. caerulescens*. Although the shoot to root cadmium concentration ratio is above unity in all the accessions tested thus far, some accessions exhibit cadmium uptake rates that are even lower than in *T. arvense* (Assunção *et al.,* 2003). Zha *et al.* (2004), studying a cross between a high-cadmium-accumulating and a low-cadmium-accumulating accession, have shown that there was little or no maternal inheritance of cadmium accumulation and their results suggest that the alleles from the low-cadmium-accumulating accession are partially dominant over those from the high-cadmium-accumulating accession. Zha *et al.* (2004) also suggested that cadmium accumulation and cadmium tolerance segregated independently in an intra-specific *T. caerulescens* F2 cross between plants with contrasting tolerance and accumulation phenotypes. Analysing inter-specific crosses, Bert *et al.* (2003) suggested that cadmium accumulation and cadmium tolerance are also under independent genetic control in *A. halleri*, as also shown for zinc accumulation and zinc tolerance in this species (Macnair *et al.,* 1999). Bert *et al.* (2003) suggest that cadmium accumulation might be a recessive character in *A. halleri*.

A number of investigations have focussed on the co-accumulation of different heavy metals. Bert *et al.* (2003) found that cadmium and zinc accumulation in aerial parts co-segregated in a back-cross between the

hyperaccumulator *A. halleri*, and the non-hyperaccumulator congener *A. lyrata* ssp. *petraea*. This shows that cadmium and zinc uptake are genetically correlated, suggesting that the metals are taken up (partly, at least) by the same transporter(s) or that their transporters, when different, are controlled by common regulators. In case of *T. caerulescens*, Zha *et al.* (2004) proposed that there is a root uptake system with a high affinity for zinc but low affinity for cadmium in the low-cadmium-accumulating ecotype, and that, in addition to such a root uptake system, there is a system with a high affinity for cadmium but low affinity for zinc in the high-cadmium-accumulating ecotype. The latter system is likely to be largely responsible for cadmium uptake in the high-cadmium-accumulating ecotype, but it also contributes to its zinc uptake, particularly at low environmental Cd concentrations.

Outline of the thesis

In the last decade heavy metal hyperaccumulator plants have been increasingly studied, mainly because of their potential use in phytoremediation. *Thlaspi caerulescens* is an attractive model hyperaccumulator plant, because it accommodates a high intra-specific variation in the degrees and metal-specificity patterns of tolerance and accumulation. This thesis further explores the natural genetic and intra-specific variability of the heavy metal hyperaccumulation trait in *Thlaspi caerulescens*. Natural genetic variation was analysed using a F_2 population (Chapter 2) and its derived F_3 population (Chapter 3), specifically developed for this thesis and genotyped with amplified fragment length polymorphism (AFLP) and other PCR based markers.

Chapter 2 further describes the construction of a linkage map and the mapping of QTL for Cd and Zn hyperaccumulation. In Chapter 3, the co-linearity patterns of the *Thlaspi caerulescens* map and the *Arabidopsis thaliana* physical map are presented. In addition, QTL for Cd, Zn and Ni accumulation in shoot and root, respectively, have been mapped. Chapter 4 reports on the phylogeny analysis of 25 European populations of *Thlaspi caerulescens*. The main results of the experiments described in the previous chapters are discussed in Chapter 5, with perspectives for future research.

References

Assunção AGL, Da Costa Martins P, De Folter S, Vooijs R, Schat H, Aarts MGM (2001) Elevated expression of metal transporter genes in three accessions of the metal hyperaccumulator *Thlaspi caerulescens*. Plant, Cell and Environment 24**,** 217-226

Assunção AGL, Schat H, Aarts MGM (2003a) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. New Phytologist 159, 351-360

Assunção AGL, Ten Bookum WM, Nelissen HJM, Vooijs R, Schat H, Ernst WHO (2003b) A co-segregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator *Thlaspi caerulescens*. New Phytologist 159: 383-390

Assunção AGL, Ten Bookum WM, Nelissen HJM, Vooijs R, Schat H, Ernst WHO (2003c) Differential metal-specific tolerance and accumulation patters among *Thlaspi caerulescens* populations originating from different soil types. New Phytologist 159: 411-419

Assunção AGL, Pieper B, Vromans J, Lindhout P, Aarts MGM, Schat H (2006) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. New Phytologist 170, 21-32

Baker AJM, Brooks R (1989) Terrestrial and higher plants which hyperaccumulate metallic elements – a review of their distribution, ecology and phytochemistry. Biorecovery 1, 81-126

Baker AJM, Reeves RD, McGrath SP (1991) *In situ* de-contamination of heavy metal polluted soils using crops of metal accumulating

plants – a feasibility study. In: Hinchee, RE, Olfenbuttel RF, eds. Situ Bioreclamation. Stoneham, MA, USA: Butterworth-Heinemann, pp 539-544

Baker AJM, McGrath SP, Reeves DR, Smith JAC (2000) Metal hyperaccumulators plants: a review of the ecology and physiology of a biological resource for phytoremediation of metal-polluted soils. In: Terry N, Banuelos G, editors. Phytoremediation of contaminated soil and water. Boca Raton, FL: CRC Press LLC: pp 85-107

Becher M, Talke IN, Krall L, Kramer U (2004) Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. Plant Journal 37, 251-268

Bernard C, Roosens N, Czernic P, Lebrun M, Verbruggen N (2004) A novel CPx-ATPase from the cadmium hyperaccumulator *Thlaspi caerulescens*. FEBS letters 569: 140-148

Bert V, Meerts P, Saumitou-Laprade P, Salis P, Gruber W, Verbruggen N (2003) Genetic basis of Cd tolerance and hyperaccumulation in *Arabidopsis halleri*. Plant and Soil 249**,** 9-18

Bovet L, Kammer PM, Meylan-Bettex M, Guadagnuolo R, Matera V (2006) Cadmium accumulation capacities of *Arabis alpina* under environmental conditions. Environmental and Experimental Botany 57, 80-88

Boyd RS (1998) Hyperaccumulation as a plant defensive strategy. In: Brooks RR, ed. Plants that hyperaccumulate heavy metals. Wallingford, UK: CAB International, 181-201

Boyd RS, Martens SN (1992) The raison d'être for metal hyperaccumulation by plants. In: Baker AJM, Proctor J, Reeves RD, eds. The vegetation of Ultramafic (Serpentine) soils. Andover, UK: Intercept, 279-289

Boyd RS, Davis MA, Wall MA, Balkwill K (2002) Nickel defends the South African hyperaccumulator *Senecio coronatus* (Asteraceae) against *Helix aspersa* (Mollusca: Pulmonidae). Chemoecology 12**,** 91-97

Boyd RS (2007) The defense hypothesis of elemental hyperaccumulation: status, challenges and new directions. Plant and Soil 293, 153-176

Brooks RR, lee J, Reves RD, Jaffre T (1977) Detection of nickeliferous rocks by analysis of herbarium specimens of indicator plants. Journal of Geochemical Exploration 7, 49-57

Courbot M, Willems G, Motte P, Arvidsson S, Roosens N, Saumitou-Laprade P, Verbruggen N (2007) A major quantitative trait locus for cadmium tolerance in *Arabidopsis halleri* colocalizes with HMA4, a gene encoding a heavy metal ATPase. Plant Physiology 144, 1052-1065

Dechamps C, Lefèbvre C, Noret N, Meerts P (2007) Reaction norms of life history traits in response to zinc in *Thlaspi caerulescens* from metalliferous and nonmetalliferous sites. New Phytologist 173, 191-198

Delhaize E, Kataoka E, Hebb DM, White RG, Ryan PR (2003) Genes encoding proteins of the cation diffusion facilitator family that confer manganese tolerance. Plant Cell 15, 1131-1142

Dubois S, Cheptou PO, Petit C, Meerts P, Poncelet M, Vekemans X, Lefebvre C, Escarré J (2003) Genetic structure and mating systems of metallicolous and nonmetallicolous populations of *Thlaspi caerulescens*. New Phytologist 157, 633-641

Escarré J, Lefèbvre C, Gruber W, Leblanc M, Lepart J, Rivière Y, Delay B (2000) Zinc and cadmium hyperaccumulation by *Thlaspi caerulescens* from metalliferous and nonmetalliferous sites in the Mediterranean area: implications for phytoremediation. New Phytologist 145, 429-437

Filatov V, Dowdle J, Smirnoff N, Ford-Lloyd B, Newbury HJ, Macnair MR (2006) Comparison of gene expression in segregating families identifies genes and genomic regions involved in a novel adaptation, zinc hyperaccumulation. Molecular Ecology 15, 3045-3059

Filatov V, Dowdle J, Smirnoff N, Ford-Lloyd B, Newbury HJ, Macnair MR (2007) A quantitative trait loci analysis of Zinc hyperaccumulation in *Arabidopsis halleri*. New Phytologist 174, 580-590

Frérot H, Petit C, Lefèbvre C, Gruber W, Collin C, Escarré J (2003) Zinc and cadmium accumulation in controlled crosses between metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* (Brassicaceae). New Phytologist 157**,** 643-648

Frey B, keller C, Zierold K, Schulin R (2000) Distribution of Zn in functionally different leaf epidermal cells in the hyperaccumulator *Thlaspi caerulescens.* Plant Cell and Environment 23, 675-687

Green LS, Rogers EE (2004) FRD3 Controls iron localization in Arabidopsis. Plant Physiology 136, 2523-2531

Guimaraes MD, Gustin JL, Salt DE (2009) Reciprocal grafting separates the roles of the root and shoot in zinc hyperaccumulation in *Thlaspi caerulescens*. New Phytologist 184, 323-329

Gygi Sp, Rochon Y, Franza BR, Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. Molecular Cell Biology 19, 1720-1730

Hajduch M, Rakwal R, Agrawal GK, Yonekura M, Pretova A (2001) High-resolution two-dimensional electrophoresis separation of proteins from metal-stressed rice (*Oryza sativa* L.) leaves: drastic reductions/fragmentation of ribulose-1,5-bisphosphate carboxylase/oxygenase and induction of stress-related proteins. Electrophoresis 22, 2824-2831

Hammond JP, Bowen HC, White PJ, Mills V, Pyke KA, Baker AJM, Whiting SN, May ST, Broadley MR (2006) A comparison of the *Thlaspi caerulescens* and *T. arvense* shoot transcriptome. New Phytologist 170, 239-260

Hanikenne M, Talke IN, Haydon MJ, Lanz C, Nolte A, Motte P, Kroymann J, Weigel D, Kramer U (2008) Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. Nature 453, 391-U44.

Hanson B, Garifullina GF, Lindblom SD, Wangeline A, Ackley A, Kramer K, Norton AP, Lawrence CB, Pilon-Smits EAH (2003) Selenium accumulation protects *Brassica juncea* from invertebrate herbivory and fungal infection, New Phytologist 159, 461-469

Hassinen VH, Tervahauta AI, Halimaa P, Plessl M, Peräniemi S, Schat H, Aarts MGM, Servomaa K, Kärenlampi SO (2006) Isolation of Zn-responsive genes from two accessions of the hyperaccumulator plant *Thlaspi caerulescens*. Planta online first

Huitson S, Macnair MR (2003) Does zinc protect the zinc hyperaccumulator *Arabidopsis halleri* from herbivory by snails? New Phytologist 159, 453-459

Hussain D, Haydon MJ, Wang Y, Wong E, Sherson SM, Young J, Camakaris J, Harper JF, Cobbett CS (2004) P-type ATPase heavy metal transporters with roles in essential zinc homeostasis in Arabidopsis. Plant Cell 16, 1327-1339

Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlet DR, Aebersold R, Hed L (2001) Integrated genomic and proteomic analyses of a systematically perturbed network. Science 292, 929-934

Ingle RA, Smith JAC, Sweetlove LJ (2005) Responses to nickel in the proteome of the hyperaccumulator plant *Alyssum lesbiacum*. Biometals 18, 627-641

Jiang RF, Ma DY, Zhao FJ, McGrath SP (2005) Cadmium hyperaccumulation protects *Thlaspi caerulescens* from leaf feeding damage by trips (*Frankliniella occidentalis*). New Phytologist 167, 805-813

Jiménez-Ambriz G, Petit C, Bourrié I, Dubois S, Olivieri I, Ronce O (2006) Life history variation in the heavy metal tolerant plant *Thlaspi caerulescens* growing in a network of contaminated and

noncontaminated sites in southern France: role of gene flow, selection and phenotypic plasticity. New Phytologist 173, 199-215

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. Ecology Letters 7, 1225-1241

Kim DY, Bovet L, Kushnir S, Noh EW, Martinoia E, Lee Y (2006) AtATM3 is involved in heavy metal resistance in Arabidopsis. Plant Physiology 140, 922-932

Kobae Y, Uemura T, Sato MH, Ohnishi M, Mimura T, Nakagawa T, Maeshima M (2004) Zinc transporter of *Arabidopsis thaliana* AtMTP1 is localized to vacuolar membranes and implicated in zinc homeostasis. Plant Cell Physiology 45, 1749-1758

Koch M, Mummenhoff K, Hurka H (1998) Systematics and evolutionary history of heavy metal tolerant *Thlaspi caerulescens*. Western Europe: evidence from genetic studies based on isozyme analysis. Biochemical Systematics and Ecology 26, 823-838

Kupper, H, Zhao FJ, McGrath SP (1999) Cellular compartimentation of zinc in the leaves of the hyperaccumulator *Thlaspi caerulescens*. Plant Physiology 199, 305-311

Lasat MM, Baker AJM, Kochian LV (1996) Physiological characterization of root Zn2+ absorption and translocation to shoots in Zn hyperaccumulator and nonaccumulator species of Thlaspi. Plant Physiology 112, 1715-1722

Ledesma A, Moral V, Villalba M, Salinas J, Rodriguez R (2006) Ca^{2+}-binding allergens from olive pollen exhibit biochemical and immunological activity when expressed in stable transgenic Arabidopsis. FEBS Journal 273**:** 4425-4434

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of polymerase chain reaction. Science 257, 967-971

Long XX, Yang XE, Ni WZ (2002) Current status and perspective on phytoremediation of heavy metal polluted soils. Journal of Applied Ecology 13, 757-62

Macnair MR (1983) The genetic control of copper tolerance in the yellow monkey flower, *Mimulus guttatus*. Heredity 50, 283-293

Macnair MR (1993) The genetics of metal tolerance in vascular plants. New Phytologist 124, 541-559

Macnair MR, Gardner M (1999) The evolution of edaphic endemics. In: Howard DJ, Berlocher SH, eds. Endless forms: species and speciation. Oxford, UK: Oxford University Press, pp 157-171

Macnair MR, Tilstone GH, Smith SE (2000) The genetics of metal tolerance and accumulation in higher plants. In: Terry N, Banuelos G, editors. Phytoremediation of contaminated soil and water. Boca Raton, FL: CRC Press LLC: pp 235-50

Macnair MR (2003) The hyperaccumulation of metals by plants. Advances in Botanical Research 40, 63-105

Mandaokar A, Kumar VD, Amway M, Browse J (2003) Micro-array and differential display identify genes involved in jasmonate-dependant anther development. Plant Molecular Biology 52, 775-786

McGrath SP, Zhao FJ (2003) Phytoextraction of metals and metalloids from contaminated soils. Current Opinion in Biotechnology 14, 277-282

McIntyre T (2003) Phytoremediation of heavy metals from soils. Advances in Biochemical Engineering/Biotechnology 78, 97-123

Meerts P, Van Isacker N (1997) Heavy metal tolerance and accumulation in metallicolous and non-metallicolous populations of *Thlaspi caerulescens* from continental Europe. Plant Ecology 133, 221-231

Mills RF, Francini A, Ferreira da Rocha PS, Baccarini PJ, AylettM, Krijger GC, Williams LE (2005) The plant P1B-type ATPase *AtHMA4* transports Zn and Cd and plays a role in detoxification of transition metals supplied at elevated levels. FEBS Letters 579, 783-791

Mukherjee I, Campbell NH, Ash JS, Connoly EL (2005) Expression profiling of the Arabidopsis ferric chelate reductase (FRO) gene family reveals differential regulation by iron and copper. Planta 14, 1-13

Noret N, Meerts P, Tolrà R, Poschenrieder C, Barceló J, Escarré J (2005) Palatability of *Thlaspi caerulescens* for snails: influence of zinc and glucosinolates. New Phytologist 165, 763-772

Ortiz, DF, Kreppel L, Speiser DM, Scheel G, McDonald G, Ow DW (1992) Heavy metal tolerance in the fission yeast requires an ATP-binding cassette-type vacuolar membrane transporter. EMBO Journal 11, 3491-3499

Papoyan A, Kochian LV (2004) Identification of *Thlaspi caerulescens* genes that may be involved in heavy metal hyperaccumulation and tolerance. Characterization of a novel heavy metal transporting ATPase. Plant Physiology 159, 421-430

Peer WA, Mamoudian M, Lahner B, Reeves RD, MurphyAS, Salt DE (2003) Identifying model metal hyperaccumulating plants: germplasm analysis of 20 Brassicaceae accessions from a wide geographical area. New Phytologist 159: 421-430

Peer WA, Mahmoudian M, Freeman JL, Lahner B, Richards EL, Reeves RD, Murphy AS, Salt DE (2006) Assessment of plants from the Brassicaceae family as genetic models for the study of nickel and zinc hyperaccumulation. New Phytologist 172, 248-260

Pence NS, Larsen PB, Ebbs SD, Letham DLD, Lasat MM, Garvin DF, Eide D, Kochian LV (2000) The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator *Thlaspi caerulescens*. Proceedings of the National Academy of Sciences, USA 97: 4956-4960

Piñeros MA, Kochian LV (2003) Differences in whole-cell and single-channel ion currents across the plasma membrane of mesophyll cells from two closely related Thlaspi species. Plant Physiology 131, 583-594

Plessl M, Rigola D, Hassinen V, Aarts MGM, Schat H, Ernst D (2005) Transcription profiling of the metal-hyperaccumulator *Thlaspi caerulescens* (J. & C. PRESL). Zeitschrift fur Naturforschung 60: 216-223

Pollard A, Baker AJM (1996) Quantitative genetics of zinc hyperaccumulation in *Thlaspi caerulescens*. New Phytologist 132, 113-118

Pollard AJ (2000) Metal hyperaccumulation: a model system for co-evolutionary studies. New Phytologist 146, 179-181

Pollard AJ, Powell KD, Harper FA, Smith JAC (2002) The genetic basis of hyperaccumulation in plants. Critical Reviews in Plant Science 21**,** 539-566

Poschenrieder C, Tolrà R, Barcelo J (2006) Can metals defend plants against Abiotic stress? Trends in Plant Science 11, 288-295

Reeves RD (1988) Nickel and zinc accumulation by species of *Thlaspi* L., *Cochlaria* L., and other genera of the Brassicaceae. Taxon 37, 309-318

Reeves RD (1992) New Zealand serpentines and their flora. In: Baker AJM, Proctor J, Reeves RD, eds. The vegetation of ultramafic (serpentine) soils. Andover, UK: Intercept, pp 129-137

Repetto O, Bestel-Corre G, Dumas-Gaudot E, Berta G, Gianinazzi-Pearson V, Gianinazzi S (2003) Targeted proteomics to identify cadmium-induced protein modifications in *Glomus mosseae*-inoculated pea roots. New Phytologist 157, 555-267

Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. Genetica 112, 183-198

Richau KH, Kozhevnikova AD, Seregin IV, Vooijs R, Koevoets PLM, Smith JAC, Ivanov VB, Schat H (2009) Chelation by histidine inhibits the vacuolar sequestration of nickel in roots of the hyperaccumulator *Thlaspi caerulescens*. New Phytologist 183, 106-116

Rigola, D, Fiers M, Vurro E, Aarts MGM (2006) The heavy metal hyperaccumulator *Thlaspi caerulescens* expresses many species-

specific genes as identified by comparative EST analysis. New Phytologist 170, 753-766

Robinson NJ, Procter CM, Connolly EL, Guerinot ML (1999) A ferric-chelate reductase for iron uptake from soils. Nature 397, 694-697

Roosens NH, Bernard C, Leplae R, Verbruggen N (2004) Evidence for copper homeostasis function of metallothionein (MT3) in the hyperaccumulator *Thlaspi caerulescens*. FEBS letters 577: 9-16

Roosens NH, Bernard C, Leplae R, Verbruggen N (2005) Variations in plant metallothioneins: the heavy metal hyperaccumulator *Thlaspi caerulescens* as a study case. Planta 222, 716-729

Schat H, ten Bookum WM (1992) Genetic control of copper tolerance in *Silene vulgaris*. Heredity 68, 219-229

Suzuki N, Koizumi N, Sano H (2001) Screening of cadmium-responsive genes in *Arabidopsis thaliana*. Plant, Cell and Environment 24, 1177-1188

Talke IN, Hanikenne M, Krämer U (2006) Zinc-dependent global transcriptional control, transcriptional deregulation, and higher gene copy number for genes in metal homeostasis of the hyperaccumulator Arabidopsis halleri. Plant Physiology 142, 148-167

Thiellement H, Bahrman N, Damerval C, Plomion C, Rossignol M, Santoni V, de Vienne D, Zivy M (1999) Proteomics for genetic and physiological studies in plants. Electrophoresis 20, 2013-2026

Thiellement H, Zivy M, Plomion C (2002) Combining proteomic and genetic studies in plants. Journal of Chromatography 752, 137-149

Thomine S, Wang RC, Ward JM, Crawford NM, Schroeder JI (2000) Cadmium and iron transport by members of a plant metal transporters family in *Arabidopsis* with homology to Nramp genes. Proceedings of the National Academy of Sciences, USA, 97: 4991-4996

Thomine S, Lelievre F, Debarbieux E, Schroeder JI, Barbier-Brygoo H (2003) *AtNRAMP3*, a multispecific vacuolar metal transporter involved in plant responses to iron deficiency. Plant Journal 34, 685-695

Tuomainen MH, Nunan N, Lehesranta SJ, TervahautaAI, Hassinen VH, Schat H, Koistinen KM, Auriola S, McNicol J, Karenlampi SO (2006) Multivariate analysis of protein profiles of metal hyperaccumulator *Thlaspi caerulescens* accessions. Proteomics 6, 3696-3706

Urquhart C (1971) Genetics of lead tolerance in *Festuca ovina*. Heredity 26, 19-33

van de Mortel JE, Aarts MGM (2006a) Comparative transcriptomics – model species lead the way. New Phytologist 170, 199-201

van de Mortel JE, Almar Villanueva L, Schat H, Kwekkeboom J, Coughlan S, Moerland PD, Ver Loren van Themaat E, Koornneef M, Aarts MGM (2006b) Large Expression Differences in Genes for Iron and Zinc Homeostasis, Stress Response, and Lignin Biosynthesis Distinguish Roots of Arabidopsis thaliana and the Related Metal Hyperaccumulator *Thlaspi caerulescens*. Plant Physiology 142, 1127-1147

van der Zaal BJ, Neuteboom LW, Pinas JE, Chardonnens AN, Schat H, Verkleij JA, Hooykaas PJ (1999) Overexpression of a novel Arabidopsis gene related to putative zinc-transporter genes from animals can lead to enhanced zinc resistance and accumulation. Plant Physiology 119, 1047-1055

Verbruggen N, Hermans C, Schat H (2009) Molecular mechanisms of metal hyperaccumulation in plants. New Phytologist 182, 781-781

Verret F, Gravot A, Auroy P, Leonhardt N, David P, Nussaume L, Vavasseur A, Richaud P (2004) Overexpression of *AtHMA4* enhances root-to-shoot translocation of zinc and cadmium and plant metal tolerance. FEBS Letters 576: 306-312

Weber M, Harada E, Vess C, Roepenack-Lahaye E, Clemens S (2004) Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as a potential metal hyperaccumulation factors. Plant Journal 37, 269-281

Wemmie JA, Szczypka MS, Thiele DJ, Moyle-Rowley WS (1994) Cadmium tolerance mediated by yeast AP-1 protein requires the presence of an ATP-binding cassette transporter-encoding gene, *YCF1*. Journal Biological Chemistry 269, 32592-32597

Willems G, Dräger DB, Courbot M, Godé C, Verbrugen N, Saumitou-Laprade P (2007) The genetic basis of zinc tolerance in the metallophyte *Arabidopsis halleri* ssp. *halleri* (Brassicaceae): an analysis of quantitative trait loci. Genetics 176, 659 – 674

Wu H, Li L, Yuan Y, Cheng X, Ling HQ (2005) Molecular and biochemical characterisation of the Fe(III) chelate reductase gene family in *Arabidopsis thaliana*. Plant Cell Physiology 46, 1505-1514

Zha HG, Jiang RF, Zhao FJ, Vooijs R, Schat H, Barker JHA, McGrath SP (2004) Co-segregation analysis of cadmium and zinc accumulation in *Thlaspi caerulescens* interecotypic crosses. New Phytologist 163, 299-312

Chapter 2

QTL analysis of cadmium and zinc accumulation in the heavy metal hyperaccumulator *Thlaspi caerulescens*

Deniau AX, Pieper B, Ten Bookum WM, Lindhout P, Aarts MGM and Schat H

Abstract

Thlaspi caerulescens (*Tc*; 2n=14) is a natural Zn, Cd and Ni hyperaccumulator species belonging to the Brassicaceae family. It shares 88% DNA identity in the coding regions with *Arabidopsis thaliana* (*At*) (Rigola *et al.,* 2006). Although the physiology of heavy metal (hyper) accumulation has been intensively studied, the molecular genetics are still largely unexplored. We address this topic by constructing a genetic map based on AFLP® markers and Expressed Sequence Tags (ESTs).

To establish a genetic map, an F_2 population of 129 individuals was generated from a cross between a plant from a Pb/Cd/Zn-contaminated site near La Calamine (LC), Belgium, and a plant from a comparable site near Ganges (GA), France. These two accessions show different degrees of Zn and particularly Cd accumulation. We analyzed 205 AFLP markers (of which 4 co-dominant) and 12 co-dominant EST sequences-based markers and mapped them to seven linkage groups (LGs), presumably corresponding to the seven chromosomes of *T. caerulescens*. The total length of the genetic map is 516 cM with an average density of one marker every 2.5 cM.

This map was used for Quantitative Trait Locus (QTL) mapping in the F_2. For both metal concentration in shoot and in root we mapped two QTLs for Cd and Zn each. These QTLs explain 22 to 42% of the variance of the traits measured. We found only one common locus (LG3) for Cd and Zn (concentration in root). For all QTLs GA increased the trait value except one for Zn accumulation in shoot (LG1).

Introduction

Thlaspi caerulescens is a heavy metal hyperaccumulator plant species that is able to accumulate extremely high levels of zinc (Zn) and cadmium (Cd) in its shoots (30,000 μg Zn g^{-1} dry weight (DW) and 10,000 μg Cd g^{-1} DW). For many years, it has been the subject of intensive research to better understand the mechanism of heavy metal hyperaccumulation and tolerance. *T. caerulescens* represents a potential source of genes for engineering heavy metal phytoremediation in plants (Assunção *et al.,* 2003b; Cobbett, 2003).

A major advantage of *T. caerulescens* over the other hyperaccumulator plant species is the available genetic variation for metal specificity with regard to accumulation, translocation and tolerance traits between different *T. caerulescens* accessions (Meerts and van Isacker, 1997; Escarré *et al*., 2000; Lombi *et al*., 2000; Schat *et al.,* 2000). This intraspecific variation permits a genetic analysis of these traits in segregating populations generated from intraspecific crosses, including Quantitative Trait Locus mapping (QTL) analysis (Alonso-Blanco and Koornneef, 2000). As the genus Thlaspi is closely related to the genus Arabidopsis (Koch et al., 2001), it is possible to assess the synteny between these species, such as previously done with *A. lyrata* and *A. thaliana* (Kuittinen *et al.*, 2004).

AFLP analysis has been shown to be well-suited for genotyping and map construction in many plant species. We constructed an AFLP-based comprehensive genetic linkage map based on the F_2 progeny of the inter-accession cross between La Calamine (LC) and Ganges (GA). The GA

accession combines the properties of extreme Cd accumulation and Cd tolerance, while the LC accession exhibits a relatively low Cd accumulation. The accumulation and tolerance of Zn are comparable in both accessions. Despite the similarities, the mechanisms of both Zn and Cd accumulation seem to differ between these accessions (Lombi *et al.*, 2000; Zhao *et al.,* 2002; Zha *et al.*, 2004). In addition to constructing a linkage map, we used the population to identify QTLs controlling Cd or Zn accumulation in order to further understand heavy metal hyperaccumulation, tolerance and homeostasis in *T. caerulescens*.

Material & Methods

Plant origin and crossing scheme: A cross was made between a plant grown from *T. caerulescens* J. & C. Presl seeds collected at a strongly Pb/Cd/Zn-enriched site near La Calamine, Belgium, and a plant grown from seeds collected at a similar calamine site near Ganges, France (Zhao *et al.*, 2002). The LC plant was used as mother. Based on former studies maternal inheritance was not expected (Zha *et al.*, 2004). One F_2 family, derived from a single self-pollinated F_1 plant, was sown and about 130 F_2 plants were analyzed.

Plant culture and vernalization: Plants were grown from seeds sown on moist peat. Three-week-old seedlings were transferred to 1-L polyethylene pots (two seedlings per pot) filled with modified half-strength Hoagland's nutrient solution: 3 mM KNO_3, 2 mM $Ca(NO_3)_2$, 1 mM $NH_4H_2PO_4$, 0.5 mM $MgSO_4$, 1 µM KCl, 25 µmol H_3BO_3, 2 µM $ZnSO_4$, 2 µM $MnSO_4$, 0.1 µM $CuSO_4$, 0.1 µM $(NH_4)6Mo_7O_{24}$, 20 µM Fe(Na)ethylene-diamine-tetra-acetic acid (EDTA). The pH buffer 2-(N-morpholino)ethane-sulfonic acid (MES) was added to a final concentration of 2 mM, and the pH was set at 5.5 using KOH. The solutions were replaced twice a week. All the crossings and experiments were performed in a climate chamber (20/15°C day/night; 250 µmol of photons at plant level; 14 $h.d^{-1}$; 75% relative humidity, RH). To induce flowering, pots with five-week-old plants were transferred to a cold growth cabinet (4°C day and night; 200 $\mu mol.m^{-2}.s^{-1}$ at plant level; 12

$h.d^{-1}$; ±60% RH) where they remained for 5-6 weeks, after which they returned to the climate chamber.

Cd and Zn accumulation in parental accessions and F_2: Three-week-old seedlings from a sample of bulked seeds of the parental accessions LC and GA (49 LC plants and 56 GA plants), and 129 F_2 plants were grown in nutrient solution (two seedlings per pot) containing 2 µM $ZnSO_4$ and supplemented with 5 µM $CdSO_4$ (Assunção *et al.,* 2003a). This concentration level was found to yield the highest relative difference in Cd accumulation between LC and GA (Assunção *et al.,* 2003c). The nutrient solutions were replaced by a fresh one twice a week, and after two weeks the leaves grown during metal exposure and half of the root system were harvested. Plants survive this treatment and are able to complete their life cycle. The roots were desorbed for 30 minutes in 5 mM ice-cold $Pb(NO_3)_2$. Roots and shoots were dried overnight in a stove at 70°C, wet-ashed in a 4:1 mixture of HNO_3 (65%) and HCl (37%) in Teflon bombs at 140°C for 7 h and analyzed for Cd and Zn using flame atomic absorption spectrometry (Perkin Elmer 1100B, Perkin Elmer Nederland, Nieuwerkerk a/d IJssel, The Netherlands).

Statistical analysis: ANOVA was used to test the statistical significance of the differences in metal concentrations between the parental controls. Correlations between elemental concentrations in F_2 plants were tested using the Pearson correlation coefficient. Data were transformed logarithmically before statistical analysis to obtain homogeneity of

variances. Transgression, i.e. segregation beyond the limits of the phenotype distributions of the parental controls, was tested for by comparing the observed and expected frequencies of F_2s in the lowest and the highest 2.5%-area sections of the parental control distributions. The 2.5% area limits were calculated as the antilog of the means ± t*SD. Expected frequencies were calculated assuming that 75% of the F_2s were distributed like the parental distribution to be tested against. The probabilities of observed frequencies were considered to equal the relative frequencies of the Poisson distributions with the expected frequencies as the means.

DNA extraction: DNA was extracted from two freshly harvested leaf disks of two-week-old plants following the CTAB protocol of Qi and Lindhout (1997). The original parents were lost and by taking four individuals of each original parent accession, we expect most alleles present in the original parents to be represented in the pool. DNA was extracted and pooled from four individuals per accession.

AFLP markers: The AFLP technique was performed as described by Vos *et al.* (1995), with some minor modifications as described by Qi and Lindhout (1997), using the enzyme combinations *Eco*RI/*Mse*I (*E/M*). The number of AFLP fragments (< 100 bands) and their polymorphism rates between the parental lines GA and LC (> 20%) were used as selection criteria for the initial pre-screening of 24 *E*/*M* primer combinations (PCs) (E32M11-22 and E35M11-22; Table 1). PCs with at least 15 easily scorable polymorphic AFLP fragments were selected. The

19 *E/M* PCs to be used on the F_2 population were E32M11, E32M12, E32M13, E32M14, E32M15, E32M16, E32M17, E32M18, E32M19, E32M20, E32M21, E32M22, E35M12, E35M13, E35M14, E35M17, E35M19, E35M20 and E35M22 (Table 2).

Primers/adapters	Sequences
*Eco*RI adapter	5'-CTCGTAGACTGCGTACC-3' 3'-CTGACGCATGGTTAA-5'
E00 (universal primer)	GACTGCGTACCAATTC
*Eco*RI adapter + 1 selective nucleotide	
E01	E00 + A
*Eco*RI adapter + 2 selective nucleotides	
E32	E00 + AAC
E35	E00 + ACA
*Mse*I adapter	5'-GACGATGAGTCCTGAG-3' 3'-TACTCAGGACTCAT-5'
M00 (universal primer)	GATGAGTCCTGAGTAA
*Mse*I adapter + 0 selective nucleotide	
M00	M00
*Mse*I adapter + 2 selective nucleotides	
M12	M00 + AC
M13	M00 + AG
M14	M00 + AT
M15	M00 + CA
M16	M00 + CC
M17	M00 + CG
M18	M00 + CT
M19	M00 + GA
M20	M00 + GC
M21	M00 + GG
M22	M00 + GT

Table 1: List of AFLP primers and adapters. DNA sequences are always given in the 5' to 3' orientation unless indicated otherwise.

Primer combination	Selective nucleotides	Total bands	Polymorphic Bands	Percent polymorphism
E32M11	AAC/AA	65	19	29,23
E32M12	AAC/AC	85	25	29,41
E32M13	AAC/AG	65	19	29,23
E32M14	AAC/AT	72	19	26,39
E32M15	AAC/CA	53	16	30,19
E32M16	AAC/CC	65	21	32,31
E32M17	AAC/CG	72	16	22,22
E32M18	AAC/CT	66	17	25,76
E32M19	AAC/GA	66	17	25,76
E32M20	AAC/GC	62	19	30,65
E32M21	AAC/GG	57	16	28,07
E32M22	AAC/GT	59	20	33,90
E35M12	ACA/AC	64	18	28,13
E35M13	ACA/AG	55	16	29,09
E35M14	ACA/AT	51	17	33,33
E35M17	ACA/CG	44	17	38,64
E35M19	ACA/GA	65	15	23,08
E35M20	ACA/GC	60	17	28,33
E35M22	ACA/GT	59	19	32,20
Total		1185	343	

Table 2: Number of bands and polymorphism detected by the 19 selected AFLP primer combinations

The adapter and primer sequences employed were based on the core primer design described by Vos et al. (1995). Gel images were electronically scanned with a Licor machine (Westburg, the Netherlands) and AFLP markers were dominantly scored using Quantar-Pro software (Keygene, Wageningen, the Netherlands). Each polymorphic AFLP band was identified by a code referring to the PC, followed by the estimated size of the DNA fragments in nucleotides (e.g. E32M15-143.5) with reference to the SequaMark 10-bp DNA ladder (Research Genetics,

Huntsville, Ala., USA). In order to score codominant markers, Quantar-Pro software was used. This program can detect codominant markers based on intensity differences of the corresponding AFLP bands. For each marker detected with the Quantar-Pro program, the reliability was checked by comparison with the marker profiles of the parent accessions, and the co-dominant scoring of the mapping population was checked by analysis of the monogenic segregation ratios.

CAPS and Indels: The sequences of *T. caerulescens* ESTs (Rigola *et al.*, 2006) of which the Arabidopsis orthologues are evenly distributed over the *At*-genome, have been used to develop a set of 12 co-dominant genetic markers (Table 3). Each EST sequence was compared to the *At*-genome using BLASTN (Altschul *et al.*, 1990) and intron flanking primers were designed. PCR-amplified fragments for LC and GA were sequenced and polymorphisms were determined using Vector NTI Suite 9™ (Invitrogen). Cleaved Amplified Polymorphic Sequence (CAPS) markers were designed if the polymorphism removed or created a restriction site. Large (>15 bp) insertions or deletions (indels) were scored on 2-3% agarose or metaphor gels, whereas acrylamide gels were used to separate small indels (<15 bp) (Table 3).

Mapping analysis: Marker segregation data were obtained by analyzing the entire F_2 population for AFLP markers using 18 selected *E/M* PCs together with the CAPS and Indel markers. Linkage analysis was performed with the computer software package JOINMAP 3.0 (Stam, 1993; Stam and van Ooijen, 1996). Kosambi's mapping function

(Kosambi, 1944) was used to convert recombination frequencies into map distances (cM). The pair-wise analysis obtained from JOINMAP was used to assign markers to linkage groups (LGs) with Log of Odds (LOD) score 4. When the combined map was compared with each of the parental maps, in general a similar marker order was found (data not shown).

EST/gene	Primers	*At* ortholog	Common name	Description	Marker type
RR6nr032	F 5'-GGATGATAAGATCGAACCTGAAGAGGC-3' R 5'-TGGTTCCATAACCACAAGAGAACTTGG-3'	none		none	Indels – 180 bp
RR1nr059	F 5'-ACCAACTCCAACTCTCTTCCTCCTCC-3' R 5'-AAGAAATCGGATGGTGTGACTGAAGC-3'	At1g09560	*GLP4/GLP5*	germin-like protein	CAPS - *Dra*I
ZNT1	F 5'-CATCGCCGATCTTCTTTGGAAATC-3' R 5'-TCGTCTCGACAGAGGTCTGATTCG-3'	At1g10970	*ZIP4*	metal transporter *ZIP4*, similar to *TcZNT1*	CAPS - *Mse*I
ZNT2	F 5'-CGTAAGACCCCAATGTTCTTCATCG-3' R 5'-TGAAGAAGCAGCCATTGTCTCTGG-3'	At1g60960	*IRT3*	metal transporter *IRT3*, similar to *TcZNT2*	CAPS – *Mse*I
RR7nr016	F 5'-TCGGATCTCCAGCAAACCGG-3' R 5'-AATAGCTTCGAGCTTGGCGTCG-3'	At1g78080	*AP2-TFRAP2,4*	AP2 domain transcription factor RAP2.4	CAPS - *Bcg*I
RR19nr015	F 5'-AAAGGCTTTTCTGCTTCAAACACTGTC-3' R 5'-TCAGGATGAGAAGAATCGATCATTGG-3'	At2g36540	*NIF*	NLI interacting factor family protein	CAPS - *Dde*I
RR11nr025	F 5'-GTGGTAACATCACTCTCCTCCGTGG-3' R 5'-AAGCATTTAGCACTCCTACTCCGGC-3'	At2g36830	*GAMMA-TIP1*	major intrinsic protein family	Indels – 15 bp
RR4nr003	F 5'-TGTTCCTCTTTACAAGGTCGGCG-3' R 5'-TCCTTGCCTCTTCTCGTACTCGAAC-3'	At3g19820	*DWARF1*	cell elongation protein	Indels – 22 bp
RR22nr089	F 5'-TCCGTTGTTGTAGTCGTCCAGGC-3' R 5'-CTTCGTCCTTGATGGGTTGAATCG-3'	At3g26310	*CytP450*	cytochrome P450 family	CAPS - *Hind*III
RR3nr075	F 5'-TCAGAGTTCAAGGAAGCGTTTAGCC-3' R 5'-CATCACGGTCCCAAGCTCCTTC-3'	At5g21274	*CAM6*	calmodulin-6	CAPS – *Eco*32I
RR25nr081	F 5'-TACCTCGAGTCTGAGAATCCTTCCAAG-3' R 5'-GGCCTAAACAAATTGTGCTAATTGGAG-3'	At5g23140	*ClpP2*	ATP-dependent Clp protease proteolytic subunit	CAPS - *Tru*I
RS26nr082	F 5'-AATCGCCGGGTACCGGAAAG-3' R 5'-TTACAACTCCAGCCCAGAGAGGAATC-3'	At5g67330	*NRAMP4*	NRAMP metal ion transporter 4	CAPS - *Bfa*I

Table 3: List of EST-based markers and their A. thaliana orthologous genes. F = Forward; R= Reverse.

QTL mapping: Potential QTLs for each trait were identified using the MAPQTL 5.0 package (van Ooijen, 2004). Kruskal-Wallis and interval mapping (IM) analyses were initially performed to find regions with potential QTL effects. Co-dominantly, otherwise dominantly, scored markers were then used in various combinations as co-factors in multiple QTL models in Multiple QTL Model analysis (MQM analysis also performed with MAPQTL). Using MAPQTL software, permutation tests calculated LOD score thresholds for each trait which was applied to declare the presence of a QTL. This corresponds to a general genome-wide significance of $P < 0.05$ for normally distributed data, as was determined by extensive simulation experiments (van Ooijen, 1999). The QTL graphs were prepared with MAPCHART (Voorrips, 2002).

Results

Characterization of Zn and Cd accumulation in LC, GA and the F_2 progeny: The ranges of shoot and root Zn concentrations for the GA accession (1.5-5.1 µmol g^{-1} DW and 0.5-2.6 µmol g^{-1}DW, respectively, *n=56*) were broadly overlapping with those for the LC accession (2.1-10.0 µmol g^{-1} DW and 0.3-1.6 µmol g^{-1} DW respectively; *n=49*) (Fig. 1).

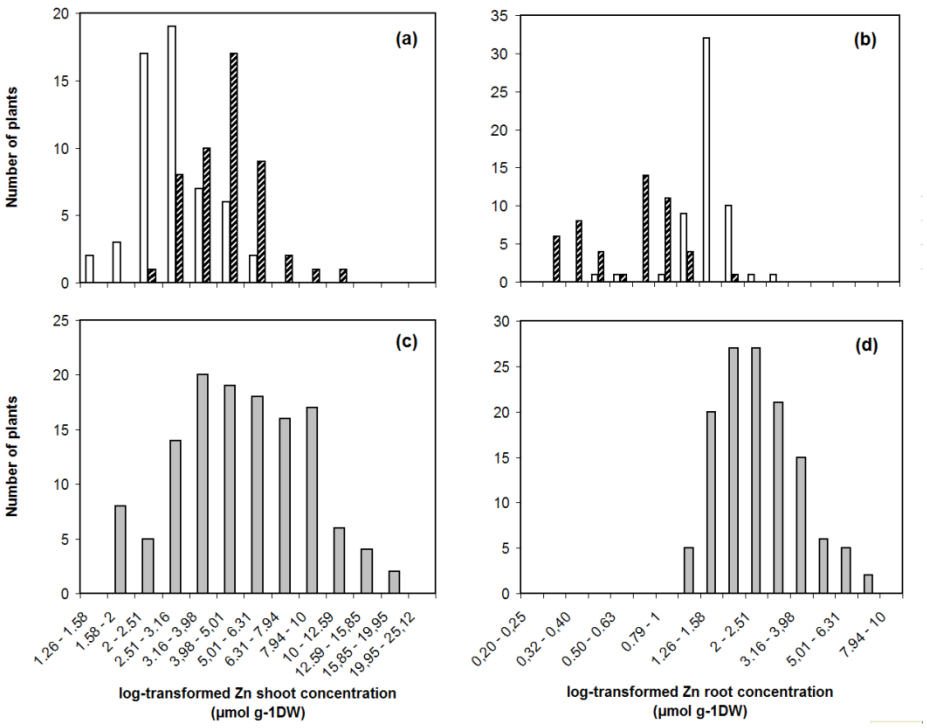


Figure 1: Frequency distributions of log-transformed Zn shoot concentration (a), root concentration (b) in the LC (dashed bars) and GA (white bars) parental controls, and in the F_2 (grey bars) from the cross LC x GA; (c) and (d) respectively. *Thlaspi caerulescens* plants were grown hydroponically with 5 µmol $CdSO_4$ and 2 µmol $ZnSO_4$.

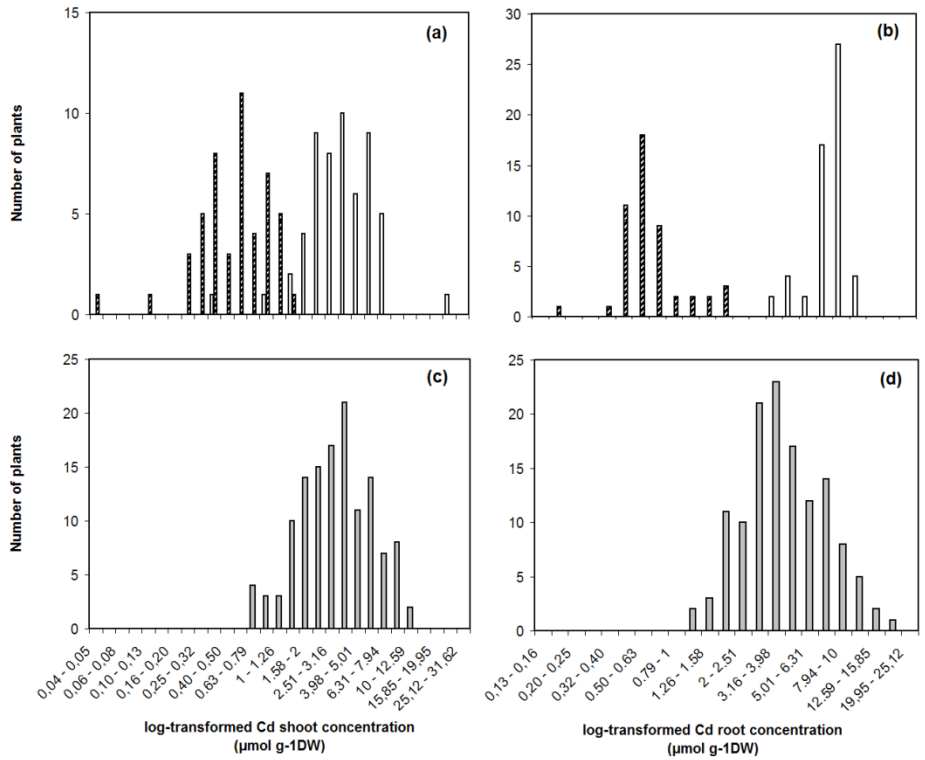


Figure 2: Frequency distributions of log-transformed Cd shoot concentration (a), root concentration (b) in the LC (dashed bars) and GA (white bars) parental controls, and in the F_2 (grey bars) from the cross LC x GA; (c) and (d) respectively. *Thlaspi caerulescens* plants were grown hydroponically with 5 µmol $CdSO_4$ and 2 µmol $ZnSO_4$

We observed only little overlap for Cd concentration in shoot (Fig. 2; 0.4-23.7 µmol g^{-1} DW for GA and 0.1-1.3 µmol g^{-1} DW for LC) and no overlap for root Cd concentrations (GA (3.2-12.4 µmol g^{-1} DW), LC (0.2-1.8 µmol g^{-1} DW)). On average, shoot and root Cd concentrations were 6.7-fold ($P < 0.001$) and 12.1-fold ($P < 0.001$) higher in GA than in LC. The Zn concentrations were 0.6-fold ($P < 0.001$) and 2.3-fold higher ($P < 0.001$), respectively (Table 4).

Trait	Population / accession	Median (in µmol.g^{-1} DW)	95-% confidence intervals (in µmol.g^{-1} DW)
ZnS	LC	4.24	3.87 - 4.65
	GA	2.72	2.51 - 2.94
	F_2	4.93	4.48 - 5.43
ZnR	LC	1.18	1.13 - 1.23
	GA	2.71	2.08 - 3.52
	F_2	2.29	2.12 - 2.46
CdS	LC	0.51	0.44 - 0.60
	GA	3.41	2.94 - 3.95
	F_2	3.80	3.40 - 4.24
CdR	LC	0.62	0.55 - 0.71
	GA	7.52	6.98 - 8.09
	F_2	3.95	3.57 - 4.37

Table 4: Median trait levels and 95-% confidence intervals for LC, GA and F_2s. ZnS = Zinc concentration in shoot, ZnR = Zinc concentration in root, CdS = Cd concentration in shoot, CdR = concentration in root.

The set of 129 F_2 plants clearly segregated for Cd and Zn shoot and root concentrations (Fig. 1 and 2). The frequency distribution of Zn concentration in shoot showed significant transgression beyond the higher limits ($P < 0.001$) of the distribution of the parental accessions. Also, there was a clear transgression towards high levels of root Zn concentration, beyond the upper limit of the GA distribution ($P < 0.001$). However, there was no significant transgression beyond the upper and lower limits of the parental distributions for any Cd-related trait. Cd and Zn concentrations were both significantly ($P < 0.01$), but far from strictly, correlated in shoot and in root (Fig. 3).

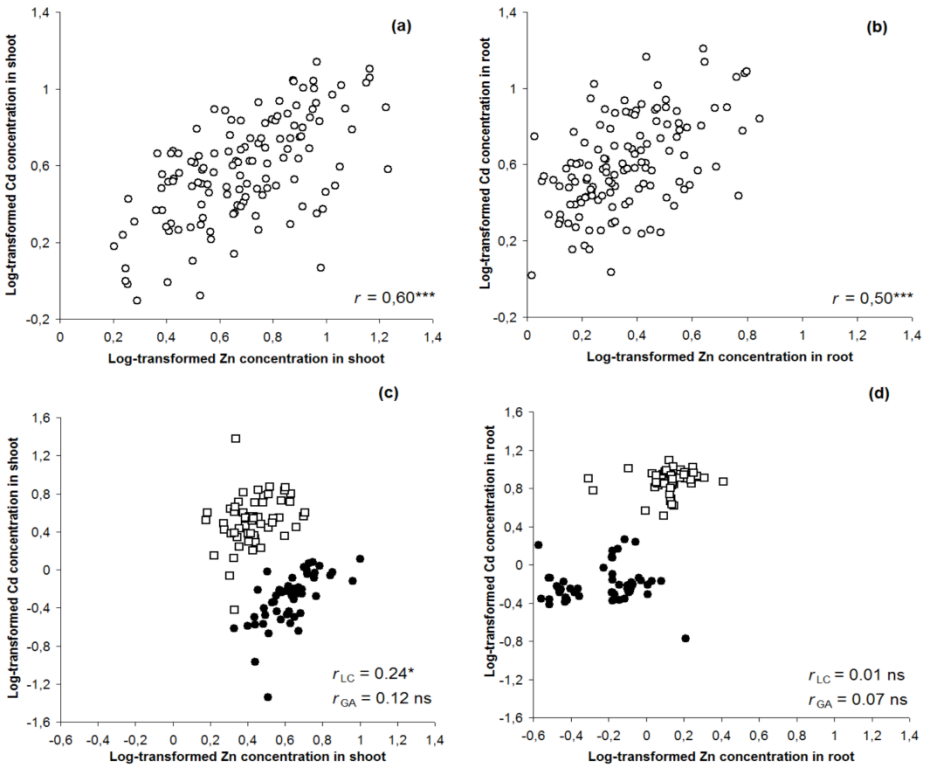


Figure 3: Pearson correlation plot of log-transformed Cd/Zn concentrations in shoot (a) and in root (b) in the F2 (open circles) and in the parental controls (LC, closed circles; GA, open squares) in shoot (c) and in root (d). Significance of the correlation: ns, not significant; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Genetic linkage map: Based on a screening for the highest polymorphism rates between both parental accessions, 19 *E*/*M* PCs were chosen. The size of the fragments generated ranged from 80 to 600 bp, with most of the fragments smaller than 500 bp. The average number of well-amplified bands varied from 44 to 85, with an average of 62.37 bands, and the average polymorphism rate between LC and GA was 29.26%. Each selected PC generated on average 18.05 polymorphisms, ranging from 15 (E35M19) to 25 (E32M12). A total of 343 polymorphic

AFLP bands were generated (157 LC-specific and 182 GA-specific) and 4 co-dominant AFLP markers were scored. In addition, 9 CAPS and 3 Indels markers were scored in the F_2 mapping population. Some of these co-dominant markers correspond to genes known to be involved in metal homeostasis, such as *TcZNT1 (*Pence *et al.*, 2000, Assunção *et al.*, 2001), *TcZNT2* (Assunção *et al*,. 2001) and the *T. caerulescens* orthologue of *AtNRAMP4* (Thomine *et al.,* 2000). A linkage map was constructed using both dominant and co-dominant markers. Of the total of 355 markers, 353 were assigned to seven LGs (LOD score 4.0), while only two dominant AFLP markers could not be fitted to any of these LGs. Markers that caused a large increase of the χ^2 value for the map of the individual chromosomes were removed from the data set. In addition when more than three AFLP markers co-segregated, the ones with the least missing data were maintained in the data set. Finally, 205 markers (189 AFLP and 16 co-dominant markers) were used to calculate an integrated map, combining the two parental AFLP maps by using the 16 co-dominant markers (Fig. 4). This map consisted of seven LGs, corresponding to the number of haploid chromosomes, each with total lengths between 57 and 89 cM. The seven LGs were labeled 1-7 in arbitrary order of grouping from JOINMAP 3.0. The total map length was 516 cM with an average of one marker per 2.5cM and a largest genetic distance of 16 cM. The marker density per cM was not significantly different between LGs. In each of the LGs, there was a cluster of linked markers, probably representing the centromeric regions. Most of the loci (75%) showed genotype ratios as expected for a segregating F_2 population (3:1). For 32 markers this ratio was

significantly distorted (P<0.05), with a cluster of three loci mapping on LG4 between position 0 and 2 cM. For these loci the GA allele was overrepresented. Two other clusters of five and eight loci mapped on LG6 between position 16 and 25 cM and 36 and 44 cM, respectively, had the LC alleles in excess. The remaining markers with a disturbed segregation were not linked to any other markers with distorted ratios.

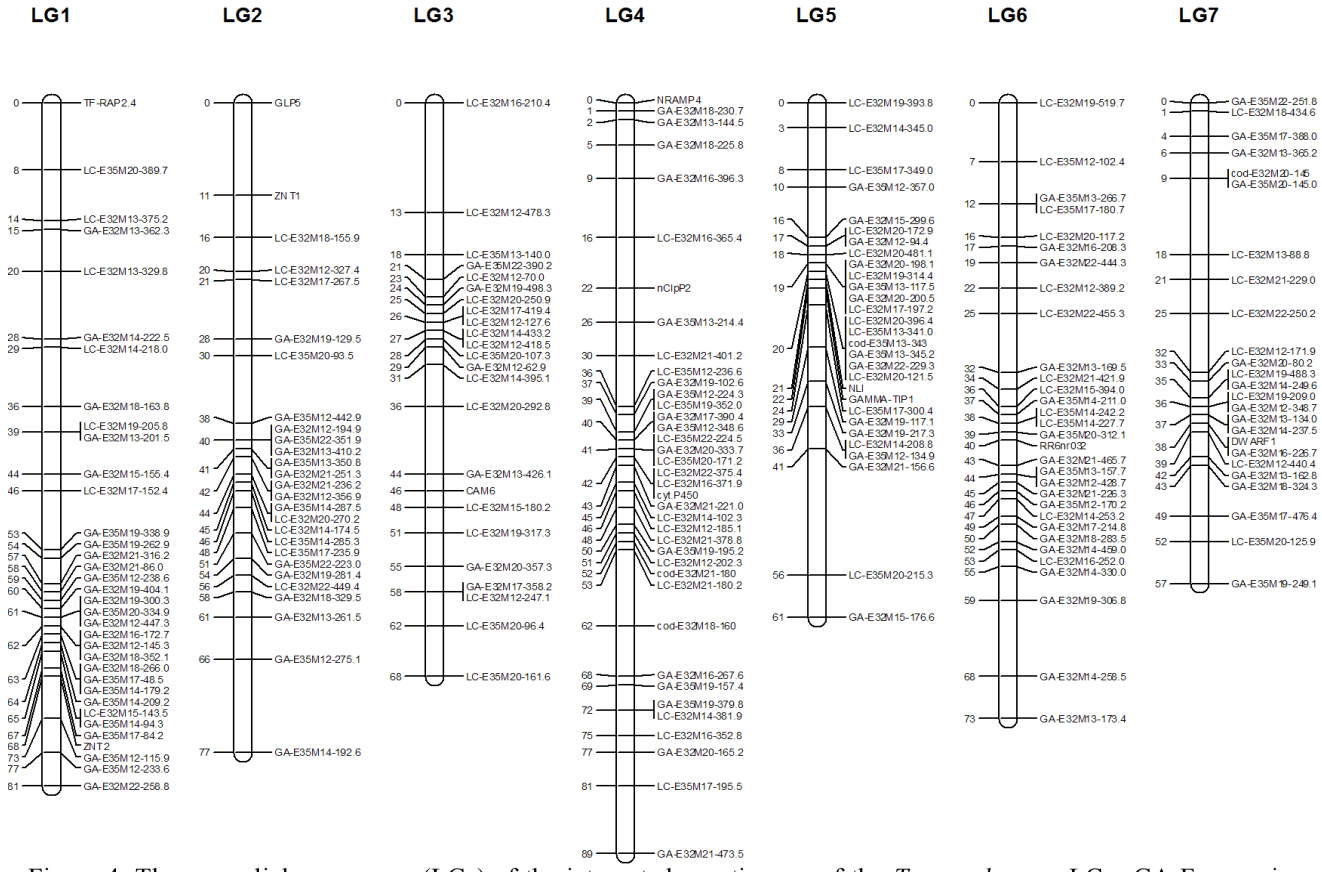


Figure 4: The seven linkage groups (LGs) of the integrated genetic map of the *T. caerulescens* LC x GA F_2 mapping population. The genetic distances are given in cM.

QTL mapping: By combining molecular marker data with data on the Cd or Zn concentration in shoot and in root, QTL mapping was performed for the four traits analyzed. Two significant QTLs were found for each trait. The explained variance per locus for each trait was not more than 24% as shown in Table 5. The location of all the significant QTLs and their 3-LOD support interval are presented in Fig. 5.

Trait	Description	GW LOD threshold	QTL	LG	Interval	Interval length (cM)	MML	LOD	%PVE	Trait enhancing allele	Trait mean (in µmol.g^{-1} DW)		
											GA/GA	GA/LC	LC/LC
ZnS	Zinc concentration in shoot	2.4	ZnS1	1	GA-E32M21-316.2 - GA-E32M19-404.1	3.5	GA-E32M21-316.2	6.2	17.6	LC	nc	nc	nc
			ZnS2	7	LC-E32M21-229.0 - LC-E32M12-171.9	10.5	LC-E32M22-250.2	8.9	24.0	GA	nc	nc	nc
ZnR	Zinc concentration in root	2.8	ZnR1	3	GA-E32M13-426.1 - LC-E32M15-180.2	4.1	CAM6	4.4	11.8	GA	3,35	2,35	1,96
			ZnR2	6	GA-E32M16-208.3 - GA-E32M22-444.3	2.3	LC-E32M20-117.2	3.7	10.3	GA	nc	nc	nc
CdS	Cadmium concentration in shoot	2.7	CdS1	3	LC-E32M15-180.2 - GA-E32M20-357.3	6.8	LC-E32M19-317.3	4.0	11.2	GA	nc	nc	nc
			CdS2	7	GA-E32M13-134.0 - GA-E32M14-237.5	0.4	GA-E32M14-237.5	4.4	10.6	GA	nc	nc	nc
CdR	Cadmium concentration in root	2.6	CdR1	3	LC-E32M20-292.8 - LC-E32M15-180.2	7.2	CAM6	4.1	10.8	GA	6,95	4,98	2,98
			CdR2	6	LC-E35M12-102.4 - GA-E35M13-266.7	5.4	LC-E35M12-102.4	5.9	16.9	GA	nc	nc	nc

Table 5: QTLs detected for Zn and Cd accumulation in shoot and root, respectively. QTLs are named according to traitt abbreviations and followed by a number to distinguish QTLs mapping affecting the same trait. LG = linkage group; MML = most closely associated Molecular Marker Loci; LOD = maximum LOD score; GW LOD Threshold = Genome Wide log-likelihood threshold determined by MapQTL5 permutation test; % PVE = percent phenotypic variation explained, nc = not calculated.

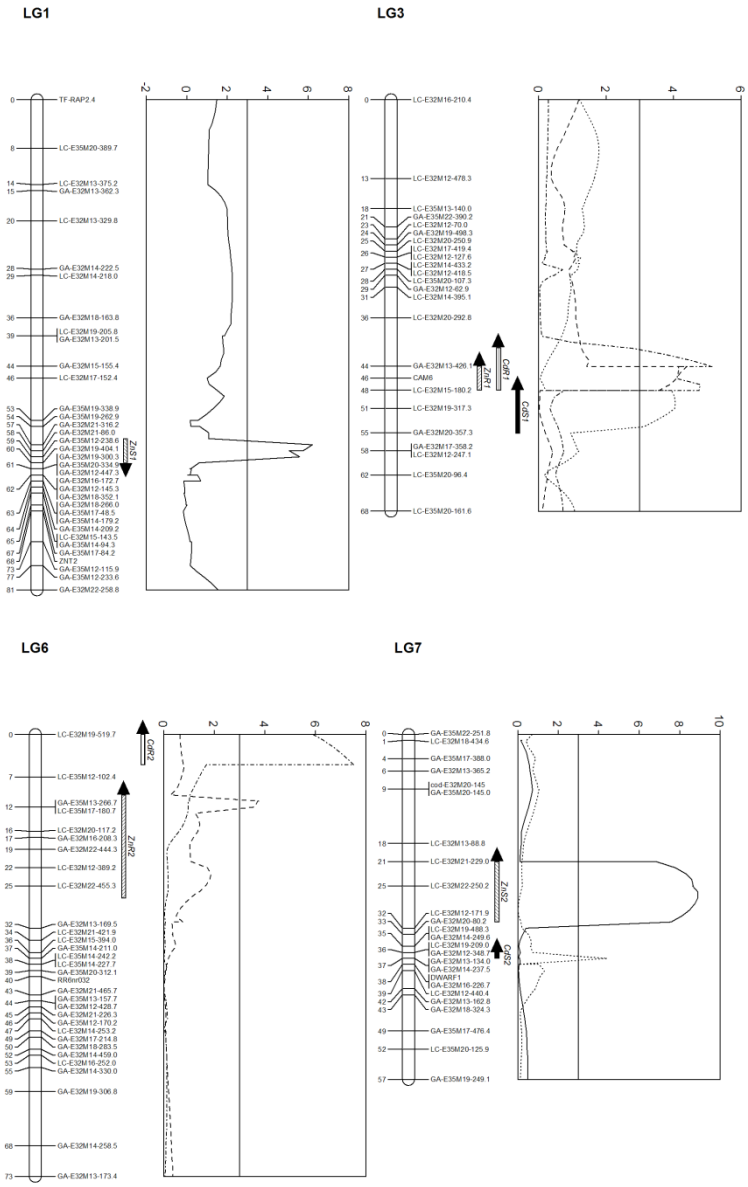


Figure 5: Location of all the significant QTLs for Zn and Cd accumulation in shoot and in root, respectively, supported by their respective genome-wide LOD support interval. The positions and origins of QTLs are indicated by upward (GA) and downward (LC) arrows. The LOD profile of each phenotype is shown on the left of the LGs (Zn accumulation in shoot, dark line; Zn accumulation in root, dashed line; Cd accumulation in shoot, dotted line; Cd accumulation in root, dashed-dotted line).

For Zn concentration in shoot we identified one QTL on LG1 (ZnS1), and another one on LG7 (ZnS2). We also found two QTLs for Zn concentration in root (ZnR) located on LG3 (ZnR1) and on LG6 (ZnR2). There were two significant QTLs for Cd concentration in shoot (CdS), on LG3 (CdS1) and LG7 (CdS2) respectively, and two for Cd concentration in root (CdR) located on LG3 (CdR1) and on LG6 (CdR2). The only trait-enhancing allele originating from LC is ZnS1. For all the other QTLs, the trait-enhancing alleles originated from GA. ZnR1 and CdR1 on LG3 are co-locating and closely linked to CdS1. CdR2 and ZnR2 on LG6, as well as ZnS2 and CdS2 on LG7 are close together but they did not map in the same interval.

Due to the lack of sufficient closely linked co-dominant markers for the identified QTLs, we were unable to establish any epistatic interactions between QTLs.

Discussion

Segregation for Zn and Cd hyperaccumulation traits: We observed that GA accumulates more Cd (7.52 µmol g^{-1}DW in GA, 0.62 µmol g^{-1}DW in LC) and Zn (2.71 µmol g^{-1}DW in GA, 1.18 µmol g^{-1}DW in LC) in root and more Cd (3.41 µmol g^{-1}DW in GA, 0.51 µmol g^{-1}DW in LC) in shoot than LC. The Zn concentration in shoot is higher in LC (2.72 µmol g^{-1}DW in GA, 4.24 µmol g^{-1}DW in LC). Previously Zha *et al.* (2004) compared shoot Zn and Cd accumulation in GA and Prayon (PR), Zn or Cd accumulation is very similar in of the PR accession compared to the LC accession (H. Schat unpublished data). At a 5-fold excess of Zn over Cd in the nutrient solution, they found similar shoot Zn accumulation in GA and PR, but at equimolar exposure PR accumulated over two times more Zn in the shoot than GA. This compares well with our observations. Overall, they found two to three times more shoot Cd accumulation in GA than in PR, which is less different than in our study (7-fold). The discrepancy may be due to the fact that we used a 2.5-fold excess of Cd over Zn, whereas Zha *et al.* (2004) supplied Zn and Cd at equimolar concentrations, or supplied Zn at a 5-fold higher concentration than Cd. Zha *et al.* (2004) proposed that two uptake systems with distinctive affinities for Cd and Zn are differently expressed in the two ecotypes. Their physiological evidence suggests that there is an uptake system with high affinity for Zn and lower affinity for Cd, which is predominant in PR, and one with high affinity for Cd but lower affinity for Zn, which is predominant or exclusive in GA. The latter system is likely to be responsible for Cd uptake in GA but also contributes to Zn

accumulation. When the Cd : Zn molar ratio in the nutrient solution is high (e.g. 2.5 in our experiment and 1.0 in Zha *et al*.'s experiment), LC and PR accumulate more Zn than GA.

The frequency distributions of the F_2 population for Cd or Zn concentration in shoot and root are not bi-modal suggesting that Cd and Zn accumulation are quantitatively inherited traits. Similar findings were described in two other studies (Assunção *et al.,* 2003b; Zha *et al.,* 2004). The Zn and Cd accumulation in shoot and in root were significantly correlated, which was also found for the Zn concentration in the study of Zha *et al.* (2004). In both cases the correlations are far from strict however, which could result from the contribution of less metal-specific mechanisms with different affinities for both metals such as suggested by Zha *et al.* (2004).

We observed a significant transgression beyond the higher limit ($P < 0.001$) of the Zn parental distributions indicating that the trait-enhancing alleles from both parents at the different QTLs add both to the phenotype. Alternative explanations for transgression are overdominance and epistasis. Overdominance is unlikely both for Cd and Zn QTLs. In all cases of inter-ecotypic Thlaspi crosses reported so far, F_1 plants showed accumulation rates that were intermediate between the parents, or dominance was directed towards low accumulation (Frérot *et al.,* 2003, Frérot *et al.,* 2005). Epistasis cannot be tested for because we were not able to distinguish between homozygote and heterozygote genotypes for both loci in a number of individuals sufficient for statistical analysis. By contrast, there was no significant transgression for Cd accumulation, which is again in agreement with Zha's results. This could be due to a

low expression of the high affinity Cd uptake system in LC, such as postulated by Zha *et al.* (2004).

The identification of two QTLs for Zn accumulation in shoot with trait enhancing alleles originating from both parents and the significant transgression found for Zn accumulation in shoot is in agreement with the hypothesis of two systems for Zn accumulation as suggested by Zha *et al.* (2004). This does not seem to be the case for Cd accumulation for which all trait enhancing alleles in the F_2 originate from GA.

Clustering of markers and map reliability: While the overall distribution of AFLP markers appeared satisfactory, some of the *E/M* clustered in distinct regions with higher marker densities. This high degree of clustering of these markers at central regions is a notable feature in many species such as barley (Qi and Lindhout, 1996), tomato (Tanksley *et al.*, 1992) or wheat (Chao *et al*., 1989) and Arabidopsis (Alonso Blanco *et al.*, 1998). Genetic maps are calculated from the recombination rates between loci as a result of chromosome crossovers at meiosis. So, assuming that markers are randomly distributed over the length of the chromosome, centromeric suppression of recombination is most likely the main reason for the clustering of markers (Tanksley *et al*., 1992; Frary *et al*., 1996). This enables us to locate putative centromers of the *T. caerulescens* chromosomes. We observed three clusters of skewed markers (LG4 and LG6) which show distorted segregation instead of the expected F_2 segregation ratios. Other mapping populations have also been reported to exhibit loci showing transmission

distortion (Boivin *et al.*, 2004, Kuittinen *et al.*, 2004, Yogeeswaran *et al.*, 2005).

The current map *T. caerulescens* compares very well to the genetic map made for another population using a common parent (e.g. Lellingen (LE) x LC in Assunção *et al.*, 2006), for which a similar total map length and comparable size of the corresponding LGs was observed. A comparison based on AFLP markers with LC specific band showed the same map position in both maps. The low number of common markers did not allow an integration of these two genetic maps. In the few cases where two markers with homology to Arabidopsis genes were mapped closely together on the same LG, they are located on the same Arabidopsis chromosome (e.g. LG1, LG2, LG4 and LG5), suggesting that chromosomal synteny might be conserved. The small number of orthologous markers does not enable us to make further assumptions on the whole genome synteny and to establish possible chromosomal rearrangements as was done for another zinc hyperaccumulator *Arabidopsis halleri* (G. Willems and P. Saumitou-Laprade, personal communation) and the related *A. lyrata* (Kuittinen *et al.*, 2004).

Quantitative Trait Loci analysis: The QTL analysis confirmed that all the studied traits are controlled by more than one gene, as expected from the phenotypic segregation patterns. The QTLs seemed to act in an additive way at least in part, since the plants homozygous for all trait-enhancing alleles showed the highest trait values. The QTLs detected in this study explained from 21.8% to 41.6 % (Table 5) of the total phenotypic variation of the traits measured. As we only had single plants

for each F_2 genotype, we cannot calculate the heritability of the traits in this population.

Interestingly, only one common locus for Zn and Cd concentration in the root was found on LG3. This could be the result of a common transporter or regulator with different expression in the two ecotypes. This is confirmed by the significant correlation between Zn and Cd accumulation. The QTLs CdR1 and CdS1 (LG3), CdR2 and ZnR2 (LG6), and ZnS2 and CdS2 (LG7) are mapped closely together. It cannot be excluded that these QTLs represent the same locus. However, since Zn and Cd accumulation are far from strictly correlated it is expected that metal specific loci are found.

The location of Zn transporter genes *ZNT1* (LG2) and *ZNT2* (Pence *et al*., 2000; Assunção *et al*., 2001) (LG1) or the iron deficiency induced transporter gene *NRAMP4* (Thomine *et al*., 2000) (LG4) on the genetic map did not coincide with the location of any QTL. However, CdS2 is closely linked to a *T. caerulescens* gene, which is orthologous to the Arabidopsis gene At3g19820 (DWARF1). In this region of chromosome 3 of Arabidopsis, there is the At3g20870 gene putatively encoding an uncharacterized protein, which contains the eukaryotic ZIP zinc transporter domain, but shows most homology to prokaryotic zinc transporters belonging to the ZIP family (Grass *et al.*, 2002). Cadmium is known to be transported by ZIP family proteins (Guerinot, 2000; Pence *et al*., 2000).

Assunção *et al.* (2006) mapped QTLs for Zn accumulation in a T. caerulescens F_3 cross between a plant from a non-metallicolous population (Lellingen, Luxemburg) and a plant from LC. The LE

ecotype is characterized by a relatively high Zn accumulation, compared to LC. In their study no significant QTLs for Zn accumulation in shoot were found. However, there were two significant QTLs for Zn in root, one with the trait enhancing allele derived from the LE parent and the other with the trait enhancing allele from the LC parent, comparable to the results obtained in the present study. Remarkably, the QTL with the trait enhancing allele originating from LE, named LE3, co-segregated with a co-dominant marker (indel47/48), representing At5g21274 (calmodulin 6). In our study ZnR1 co-segregates with the co-dominant marker CAM6, which represents the same gene. In addition, both QTLs are co-dominant. It is possible that the genes underlying LE3 and ZnR1 are in fact identical. The occurrence of trait enhancing alleles at identical gene loci in different geographically isolated populations might result from parallel evolution driven by selection pressure on the same locus or, alternatively, to common descent. In the latter case, natural selection might have favored a less effective allele within the LC accession, in agreement with its relatively low Zn accumulation capacity (Assunção *et al.*, 2003c). It might be argued that high rates of Zn accumulation in roots may occur as a result of a low translocation of Zn to the shoots. However, there was no evidence of such a correlation. The second QTL for Zn accumulation in root found by Assunção *et al.* (2006) was mapped to chromosome 5 (LG5 in our study), where we did not find any QTL for Zn accumulation in root. This could mean that LC and GA have similarly effective alleles at this particular locus. Taken together the results of both studies, it is evident that there are at least three loci determining the inter-accession variation in Zn accumulation in the root.

Dependent on the origin of the parents, either of these loci may or may not contribute to the segregation of the trait in inter-accession crosses.

Overall, this study has proven that AFLP markers are adequate for map construction and QTL analysis in *T. caerulescens*, which agrees with the findings of a recent study on *T. caerulescens* by Assunção *et al.* (2006). The present molecular map and the accompanying mapping population open new avenues for the identification of genes involved in Cd and Zn hyperaccumulation, which will enable us to unravel the molecular genetics of these complex traits. EST-based markers which link apparently orthologous genes from Arabidopsis and *T. caerulescens* will be used for comparative mapping and co-linearity analysis. Eventually this will facilitate the selection of candidate genes for *T. caerulescens* QTL based on DNA information from the corresponding genome region in Arabidopsis.

Acknowledgments

We are very grateful to Diana Rigola (Lab of Genetics, Wageningen University, NL) for providing the *Thlaspi caerulescens* EST data prior to publication and to Fangjie Zhao (Agriculture and Environment Division, IACR-Rothamsted, UK) for providing us with seeds of the Ganges accession. We thank Maarten Koornneef for critically reading the manuscript and Petra van den Berg, Fien Meijer-Dekens and Richard Flinkers for their technical advice and assistance. This research was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek-Genomics (NWO) Grant (050-10-166) and the European Union-PHYTAC project (QLRT-2001-00429).

References

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, Van den Bosch N, Pot J, Kuiper MTR (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. Plant Journal 14: 259-271

Alonso-Blanco C, Koornneef M (2000) Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. Trends in Plant Science 5, 22-29

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. Journal of Molecular Biology 215, 403-410

Assunção AGL, Da Costa Martins P, De Folter S, Vooijs R, Schat H, Aarts MGM (2001) Elevated expression of metal transporter genes in three accessions of the metal hyperaccumulator *Thlaspi caerulescens*. Plant, Cell and Environment 24**,** 217-226

Assunção AGL, Schat H, Aarts MGM (2003a) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. New Phytologist 159, 351-360

Assunção AGL, Ten Bookum WM, Nelissen HJM, Vooijs R, Schat H, Ernst WHO (2003b) A co-segregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator *Thlaspi caerulescens*. New Phytologist 159, 383-390

Assunção AGL, Ten Bookum WM, Nelissen HJM, Voijs R, Schat H, Ernst WHO (2003c) Differential metal-specific tolerance and

accumulation patters among *Thlaspi caerulescens* populations originating from different soil types. New Phytologist 159, 411-419

Assunção AGL, Pieper B, Vromans J, Lindhout P, Aarts MGM, Schat H (2006) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. New Phytologist 170, 21-32

Boivin K, Acarkan A, Mbulu R-S, Clarenz O, Schmidt R (2004) The Arabidopsis genome sequence as a tool for genome analysis in Brassicaceae. A comparison of the Arabidopsis and *Capsella rubella* genomes. Plant Physiology 135, 735-744

Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homologous group 7 chromosomes. Theoretical and Applied Genetics 78, 495-504

Cobbett C (2003) Heavy metals and plants model systems and hyperaccumulators. New Phytologist 159, 289-293

Escarré J, Lefèbvre C, Gruber W, Leblanc M, Lepart J, Rivière Y, Delay B (2000) Zinc and cadmium hyperaccumulation by *Thlaspi caerulescens* from metalliferous and nonmetalliferous sites in the Mediterranean area: implications for phytoremediation. New Phytologist 145, 429-437

Frary A, Presting GG, Tanksley SD (1996) Molecular mapping of the centromers of tomato chromosomes 7 and 9. Molecular Genetics and Genomics 250, 295-304

Frérot H, Petit C, Lefèbvre C, Gruber W, Collin C, Escarré J (2003) Zinc and cadmium accumulation in controlled crosses between

metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* (Brassicaceae). New Phytologist 157**,** 643-648

Frérot H, Lefèbvre C, Petit C, Collin C, Dos Santos A, Escarré J (2005) Zinc tolerance and hyperaccumulation in F_1 and F_2 offspring from intra and inter-ecotype crosses of *Thlaspi caerulescens*. New Phytologist 165, 111-119

Grass G, Wong MD, Rosen BP, Smith RL, Rensing C (2002) ZupT is a Zn(II) uptake system in *Escherichia coli.* Journal of Bacteriology 184, 864-866

Guerinot ML (2000) The ZIP family of metal transporters. Biochimica et Biophysica Acta 1465, 190-198

Koch M, Haubold B, Mitchell-Olds T (2001) Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. American Journal of Botany 88, 534-544

Kosambi D (1944) The estimation of map distances from the recombination values. Ann Eugen 12, 172-175

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M, Mitchell-Olds T, Langley C, Savolainen O (2004) Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *Arabidopsis thaliana.* Genetics 168, 1575-1584

Lombi E, Zhao FJ, Dunham SJ, McGrath SP (2000) Cadmium accumulation in populations of *Thlaspi caerulescens* and *Thlaspi goesingense*. New Phytologist 145, 11-20

Meerts P, Van Isacker N (1997) Heavy metal tolerance and accumulation in metallicolous and non-metallicolous populations of

Thlaspi caerulescens from continental Europe. Plant Ecology 133, 221-231

Pence NS, Larsen PB, Ebbs SD, Letham DLD, Lasat MM, Garvin DF, Eide D, Kochian LV (2000) The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator *Thlaspi caerulescens*. Proceedings of the National Academy of Sciences, USA 97, 4956-4960

Qi X, Lindhout P (1996) Development of AFLP markers in barley. Molecular Genetics and Genomics 254, 330-336

Rigola, D, Fiers M, Vurro E, Aarts MGM (2006) The heavy metal hyperaccumulator Thlaspi caerulescens expresses many species-specific genes as identified by comparative EST analysis. New Phytologist 170, 753-766

Schat H, Llugany M, Bernhard R (2000) Metal-specific patterns of tolerance, uptake, and translocation of heavy metals in hyperaccumulating and non-hyperaccumulating metallophytes. In: Terry N, Banuelos G, eds. Phytoremediation of contaminated soils and water. Boca Raton, FL: CRC Press, 171-188

Stam P (1993) Construction of integrated linkage maps by means of a new computer package: JOINMAP. Plant Journal 3, 739-744

Stam P, van Ooijen JW (1996) JOINMAP version 3.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen

Tanksley SD, Ganal MW, Prince JP, de-Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND (1992) High-density

molecular linkage maps of the tomato and potato genomes. Genetics 132, 1141-1160

Thomine S, Wang RC, Ward JM, Crawford NM, Schroeder JI (2000) Cadmium and iron transport by members of a plant metal transporters family in *Arabidopsis* with homology to Nramp genes. Proceedings of the National Academy of Sciences, USA, 97, 4991-4996

van Ooijen JW (1999) LOD significant thresholds for QTL analysis in experimental populations of diploid species. Heredity 89, 803-811

van Ooijen JW (2004) MapQTL® 5, software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V., Wageningen, Netherlands

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. Heredity 93, 77-78

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M (1995) AFLP®: a new technique for DNA fingerprinting. Nucleic Acids Research 23, 4407-4414

Yogeeswaran K, Frary KA, York TL, Amenta A, Lesser AH, Nasrallah JB, Tanksley SD, Nasrallah ME (2005) Comparative genome analyses of Arabidopsis spp.: Inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. Genome research 15, 505-515

Zha HG, Jiang RF, Zhao FJ, Vooijs R, Schat H, Barker JHA, McGrath SP (2004) Co-segregation analysis of cadmium and zinc

accumulation in *Thlaspi caerulescens* interecotypic crosses. New Phytologist 163, 299-312

Zhao FJ, Hamon RE, Lombi E, McLaughlin MJ, McGrath SP (2002) Characteristics of cadmium uptake in two contrasting ecotypes of the hyperaccumulator *Thlaspi caerulescens*. Journal of Experimental Botany 53, 535-543

Chapter 3

Comparative mapping and Cd/Zn/Ni accumulation QTL analysis in the heavy metal hyperaccumulator *Thlaspi caerulescens*

Deniau AX, Hakvoort HWJ, ten Bookum WM, Pieper B,

van Heusden S, Aarts MGM and Schat H

Abstract

Thlaspi caerulescens (*Tc*; 2n=14) is a natural Zn, Cd and Ni hyperaccumulator species belonging to the Brassicaceae family. It shares 88% DNA identity in the coding regions with *Arabidopsis thaliana* (*At*) (Rigola et al. 2006). Although the physiology of heavy metal (hyper)accumulation has been intensively studied, the molecular genetics are still largely unexplored. We address this topic by constructing a genetic map based on AFLP® markers and Expressed Sequence Tags (ESTs) and a comparative analysis of the *Thlaspi caerulescens* map with a unified comparative genomic framework of the Brassicaceae.

To establish a genetic map, an F_3 population of 145 individuals was generated from a cross between a plant from a Pb/Cd/Zn-contaminated site near La Calamine (LC), Belgium, and a plant from a comparable site near Ganges (GA), France. These two accessions show different degrees of Zn, Ni and particularly Cd accumulation. We analyzed 170 AFLP markers and 44 co-dominant EST sequences-based markers and mapped them to seven linkage groups (LGs). The total length of the genetic map is 521 cM with an average density of one marker every 2.45 cM.

The transition from eight chromosomes of the AK map to seven chromosomes in the *Thlaspi caerulescens* map can be explained by nine main chromosomal rearrangements (four translocations, two fusions and three inversions). For all traits except for Zn concentration in the shoot of plants under Cd exposure at least one QTL was detected. The 11 QTLs with significant effects were distributed over four LGs. Individual

QTLs accounted for 8.5-40% of the phenotypic variation in this population.

Introduction

Thlaspi caerulescens J. & S. Presl (Brassicaceae) is one of the plant models to study heavy metal hyperaccumulation. It has the ability to accumulate and tolerate >25 000 mg zinc (Zn) kg^{-1} dry weight (DW) in the shoots (Baker *et al.*, 1994; Shen *et al.*, 1997, Assunção *et al.*, 2003a). Zn hyperaccumulation is a constitutive and species-wide phenomenon in *T. caerulescens,* although there is a considerable variation in the degree of hyperaccumulation between accessions when grown under controlled conditions (Baker *et al.*, 1994; Meerts and van Isaker, 1997; Escarré *et al.*, 2000; Lombi *et al.*, 2000; Assunção *et al.*, 2003b; Roosens *et al.*, 2003). *T. caerulescens* is also able to hyperaccumulate other metals, such as nickel (Ni) or cadmium (Cd) (Baker *et al.*, 1994). When grown under the same conditions, accessions were found to vary widely in Cd accumulation (Lombi *et al.*, 2000; Assunção *et al.*, 2003b; Roosens *et al.*, 2003). In particular, accessions from southern France (Ganges (GA)) were far superior in Cd accumulation compared with those tested from other regions of Europe, including the Prayon (PR) and La Calamine (LC) accessions from Belgium, which are commonly used in physiological and molecular studies (Lombi *et al.*, 2000; Assunção *et al.*, 2001; Assunção *et al.*, 2003b; Roosens *et al.*, 2003). Recent studies using inter-ecotypic crosses of *T. caerulescens* suggest that multiple genes are involved in the accumulation of Zn (Assunção *et al.*, 2003b) and Cd (Zha *et al.*, 2004). The molecular mechanisms responsible for the

extraordinary ability to hyperaccumulate Cd in the southern French Ganges accession of *T. caerulescens* have not been elucidated.

The genetic variation in accumulation capacity among populations has also been exploited to map the loci contributing to zinc (Assunção *et al.,* 2006; Deniau *et al.,* 2006) and cadmium (Deniau *et al.,* 2006) hyperaccumulation using quantitative trait loci (QTL) analysis (Alonso-Blanco & Koornneef, 2000). However, there is no such study for Ni and simultaneous QTL analysis for Cd, Zn and Ni accumulation has not yet been performed in a single intra-specific cross. Such a simultaneous analysis is necessary to establish whether the accumulation of these metals, in so far as it segregates, is governed either by metal-specific or common genetic determinants, or both.

To pinpoint the genes responsible for the QTL it is crucial to identify the chromosomal regions corresponding to the identified genetic loci. At the moment this is not straightforward in *T. caerulescens*, since there is no physical map of the *T. caerulescens* genome. However, considering the often well-preserved genome co-linearity of members of the Brassicaceae family (Schranz *et al.* 2006), it is possible to identify an Arabidopsis genomic region co-linear with the *T. caerulescens* QTL interval by using flanking common genes as genetic markers. This allows further fine-mapping. In addition, the Arabidopsis genomic region can be searched directly for candidate genes based on a presumed shared function in metal homeostasis in both species.

In this paper, the first comparative mapping study in *Thlaspi caerulescens* is presented. In addition, QTL for Cd, Zn and Ni accumulation in shoot and root, respectively, are mapped in the intra-

specific F_3 LC x GA population derived from the F_2 population described by Deniau *et al.* (2006).

Material & Methods

Plant materials and culture conditions

A cross was made between a plant grown from *T. caerulescens* J. & C. Presl seeds collected at a strongly Pb/Cd/Zn-enriched site near La Calamine, Belgium, and a plant grown from seeds collected at a similar calamine site near St. Laurent le Minier, France, generally known as "Ganges" (Zhao *et al.,* 2002). The LC plant was used as a mother. Based on former studies maternal inheritance was not expected (Zha *et al.,* 2004). One F_2 family, derived from a single self-pollinated F_1 plant, was sown and about 130 F_2 plants were exposed to Cd and analyzed for the Cd and Zn concentrations in roots and shoots (Deniau *et al*., 2006), 100 were exposed to Zn only and analyzed for Zn, and 100 were exposed to Ni and analyzed for Ni and Zn in roots and shoots (unpublished data). F_2 plants were self-pollinated and a single-seed-descent F_3 population of 144 plants was composed. These are the combined single-seed progeny of 67 F_2 plants phenotyped for Cd and Zn accumulation (Deniau *et al*., 2006), 40 F_2 plants phenotyped for Zn accumulation only, and 37 F_2 plants phenotyped for Ni and Zn accumulation. Plants were grown from seeds sown on moist peat. Three-week-old seedlings were transferred to 1-L polyethylene pots (two seedlings per pot) filled with modified half-strength Hoagland's nutrient solution: 3 mM KNO_3, 2 mM $Ca(NO_3)_2$, 1 mM $NH_4H_2PO_4$, 0.5 mM $MgSO_4$, 1 µM KCl, 25 µmol H_3BO_3, 2 µM $ZnSO_4$, 2 µM $MnSO_4$, 0.1 µM $CuSO_4$, 0.1 µM $(NH_4)6Mo_7O_{24}$, 20 µM Fe(Na)ethylene-diamine-tetra-acetic acid (EDTA). The pH buffer 2-(N-

morpholino)ethane-sulfonic acid (MES) was added to a final concentration of 2 mM, and the pH was set at 5.5 using KOH. The solutions were replaced twice a week. All the crossings and experiments were performed in a climate chamber (20/15°C day/night; 250 µmol.sec^{-1}.m^{-2} photons flux density at plant level; 14 h.d^{-1}; 75% RH). To induce flowering, pots with five-week-old plants were transferred to a cold growth cabinet (4°C day and night; 200 µmol.m^{-2}.s^{-1} at plant level; 12 h.d^{-1}; ±60% RH) where they remained for 5-6 weeks, after which they were returned to the climate chamber. All F_3 plants were propagated by cuttings. One side shoot was removed with a sharp scalpel from the mother plant and 1 cm of the cutting was dipped in Rhizopon® B powder (Rhizopon BV, Hazerswoude, The Netherlands), which contains 0.2% 1-naphtalic acetic acid. Cuttings were planted in a plastic basket filled with a moist 1:2 mixture of sand and peat and finally covered with transparent foil for three weeks, to avoid desiccation. Rooted cuttings were adapted in hydroponic culture for one week before exposure to heavy metals.

Cd and Zn accumulation in parental accessions and F_3s

Three-week-old seedlings from a sample of bulked seeds of the parental accessions LC and GA (about 50 plants), and 145 F_3 plants were grown in nutrient solution (two seedlings or cuttings per pot) containing 2 µM $ZnSO_4$ supplemented with 5 µM $CdSO_4$ (Cd treatment) (Deniau *et al*., 2006); or 10 µM $ZnSO_4$ (Zn treatment); or 2 µM $ZnSO_4$ supplemented with 10 µM $NiSO_4$ (Ni treatment) (Assunção *et al.,* 2003a). These concentration levels were found to yield the highest relative differences in Cd, Zn and Ni accumulation between LC and GA (Assunção *et al.,*

2003c), respectively. The nutrient solutions were refreshed twice a week, and after two weeks shoots and roots were harvested. The roots were desorbed for 30 minutes in 5 mM ice-cold $Pb(NO_3)_2$. Roots and shoots were dried overnight in a stove at 70°C, wet-ashed in a 4:1 mixture of HNO_3 (65%) and HCl (37%) in Teflon bombs at 140°C for 7 h and analyzed for Zn, Cd and/or Ni using flame atomic absorption spectrometry (Perkin Elmer 1100B, Perkin Elmer Nederland, Nieuwerkerk a/d IJssel, The Netherlands).

Statistical analysis

ANOVA was used to test the statistical significance of the differences in metal concentrations between the parental controls. Correlations between elemental concentrations in F_2 and in F_3 plants were tested using the Pearson correlation coefficient. Data were transformed logarithmically before statistical analysis to obtain homogeneity of variances. We tested each of the markers for normal Mendelian segregation using a χ^2-test at a significance level of 0.05 followed by a correction according to the Bonferroni-Holm sequential method (Rice 1989).

Genomic DNA extraction

Around 100 mg of frozen *T. caerulescens* shoot tissue was ground to a fine powder and extracted in 0.4 M TRIS-HCl pH 8, 0.15 M NaCl, 60 mM EDTA, 0.27% (w/v) SDS, 1% polyvinylpyrrolidone and 5M KAc for 10 min at 65°C with frequent mixing. After incubation on ice for 20 min, ice-cold chloroform was added and the sample tube was placed on a tilt shaker for 10 min at room temperature. After centrifugation at 14 000

g for 10 min, the supernatant was mixed with an equal volume of isopropanol. The genomic DNA was pelleted by centrifugation at 14 000 *g* for 10 min and washed once with 70% (v/v) ethanol, spun at 14 000 *g* for 10 min, air-dried for 10–15 min, and resuspended in 50 µl of sterile water. The samples were afterwards treated with RNase A (final concentration 100 µg ml^{-1}) for 30 min at 37°C. Since the original parents were lost, we extracted DNA from a pool of four individuals per parental accession, expecting that these pools should contain most of the alleles present in the original parents.

AFLP markers

The AFLP technique was performed as described by Vos *et al.* (1995), with some minor modifications as described by Qi and Lindhout (1997), using the enzyme combinations *EcoR*I/*Mse*I (*E/M*) and *Pst*I/*Mse*I (*P/M*). Five *E/M* primer combinations (PCs) previously used in Deniau *et al.* (2006) were selected (E32M12, E32M13, E32M14, E32M19, E35M20). The number of AFLP fragments (< 100 bands) and their polymorphism rates between the parental lines GA and LC (> 20%), resulting in at least 15 easily scorable polymorphic AFLP fragments, were used as selection criteria for the initial pre-screening of 39 *P/M* PCs (Table 1). In addition to the five *E/M* PCs, the 11 *P/M* PCs to be used on the F_3 population were P14M48, P14M49, P14M51, P14M52, P14M53, P22M47, P22M51, P22M53, P26M49, P26M51 and P26M52. The adapter and primer sequences employed were based on the core primer design described by Vos *et al.* (1995). Gel images were electronically scanned with a Licor machine (Westburg, the Netherlands) and AFLP markers

were dominantly scored using Quantar-Pro software (Keygene, Wageningen, the Netherlands). Each polymorphic AFLP band was identified by a code referring to the PC, followed by the estimated size of the DNA fragments in nucleotides (e.g. E32M15-143) with reference to the SequaMark 10-bp DNA ladder (Research Genetics, Huntsville, Ala., USA).

CAPS and Indels

T. caerulescens EST sequences (Rigola *et al.,* 2006) corresponding to Arabidopsis orthologues that are evenly distributed over the Arabidopsis genome, have been used to develop a set of 44 co-dominant genetic markers (Table 2). putative intron flanking primers were designed. PCR-amplified fragments for LC and GA were sequenced and polymorphisms were determined using Vector NTI Suite 9™ (Invitrogen).

Primers/adapters	Sequences
*Eco*RI adapter	5'-CTCGTAGACTGCGTACC-3' 3'-CTGACGCATGGTTAA-5'
E00 (universal primer)	GACTGCGTACCAATTC
*Eco*RI + 1 selective nucleotide E01	E00 + A
*Eco*RI + 2 selective nucleotides E32	E00 + AAC
E35	E00 + ACA
*Pst*I adapter	5'-CTCGTAGACTGCGTACATGCA-3' 3'-CTGACGCATGGTTAA-5'
P00 (universal primer)	GACTGCGTACATGCAG
PstI + 0 selective nucleotide P00	P00
PstI + 2 selective nucleotides P14	P00 + AT
P22	P00 + GT
P26	P00 + TT
*Mse*I adapter	5'-GACGATGAGTCCTGAG-3' 3'-TACTCAGGACTCAT-5'
M00 (universal primer)	GATGAGTCCTGAGTAA
*Mse*I + 0 selective nucleotide M00	M00
*Mse*I + 1 selective nucleotides M03	M00 + G
*Mse*I + 2 selective nucleotides M11	M00 + AA
M12	M00 + AC
M13	M00 + AG
M14	M00 + AT
M19	M00 + GA
M20	M00 + GC
*Mse*I + 3 selective nucleotides M47	M00 + CAA
M48	M00 + CAC
M49	M00 + CAG
M51	M00 + CCA
M52	M00 + CCC
M53	M00 + CCG

Table 1: List of AFLP primers and adapters. DNA sequences are always given in the 5' to 3' orientation unless indicated otherwise

*Tc*EST /gene	*At* ortholg	Primer	Comon Name	Description	Marker type
RR9nr066	At1g05300	F- ATCACTGGGAGCATCACAAAGTCC	ZIP5	metal transporter	Indels
		R- CAAGGAGAGAGATATGGGCCATGA			
RR1nr059	At1g09560	F- ACCAACTCCAACTCTCTTCCTCCTCC	GLP5	germin-like protein	CAPS - *Dra*I
		R- AAGAAATCGGATGGTGTGACTGAAGC			
ZNT1	At1g10970	F- CGTGGTTGTGGAAAGAGGGAATG	ZIP4	Metal transporter ZIP4, similar to TcZNT1	CAPS - *Mse*I
		R- AAGAACTGGTGAAATGATAGAGCTG			
RR7nr051	At1g47128	F- GACAATCTCCGTTTCATCGACG	RD21A	cysteine proteinase	Indels
		R- CCTCCTCTGTATCGATTCCACCAT			
Tc_con 72	At1g52300	F- AAGGGAACGGGAAGTTTCGG	RPL37B	60S ribosomal protein L37	Indels
		R- CCTCATCCTTCCGGTTCCAG			
ZNT2	At1g60960	F- CGTAAGACCCCAATGTTCTTCATCG	IRT3	Metal transporter 3, similar to TcZNT2	CAPS - *Mse*I
		R- TGAAGAAGCAGCCATTGTCTCTGG			
RS20nr02	At1g65960	F- GCTGTTCAATGCACCGCTCAAG	GAD2	glutamate decarboxylase 1	Indels
		R- CCAGCAAGTGGGTTTAGACGCT			
RS19nr090	At1g70940	F- GTTTGCCATGGCGGTCAGGT	PIN3	auxin transport protein	Indels
		R- AACCAGCGTGATCGGAAGCG			

Table 2: List of EST-based markers (Rigola et al., 2006) and their A. thaliana orthologs. Primers sequences are always given in the 5' to 3' orientation. F forward, R Reverse

*Tc*EST /gene	*At* ortholg	Primer	Comon Name	Description	Marker type
RR7nr016	At1g78080	F- TCGGATCTCCAGCAAACCGG	RAP2.4	ERF/AP2 transcription factor family	CAPS - *Bcg*I
		R- AATAGCTTCGAGCTTGGCGTCG			
RR23nr019	At2g23150	F- AGAACTACGGAATAAGGAAGCTCGAGG	NRAMP3	NRAMP metal ion transporter 3	Indels
		R- GAGCTGAGTGAAGAAACACGTTGTGC			
RR19nr015	At2g36540	F- AAAGGCTTTTCTGCTTCAAACACTGTC	NIF	NLI interacting factor family protein	CAPS - *Dde*I
		R- TCAGGATGAGAAGAATCGATCATTGG			
RR11nr068	At2g46800	F- GTCTTAGGTGCATCTCTTAATTTCGGTG	MTP1A	zinc transporter	Indels
		R- CCGGCTGAGGAAGAAGGGC			
RR27nr057	At3g01910	F- TTCTACAAGAGAAATCACGGCCCC	SOX	sulfite oxidase	Indels
		R- GCAGAAACATCCCCATCCAACTC			
RR7nr084	At3g05880	F- TGACGCTGTTTGGGTATCTTCCTG	RCI2A	hydrophobic protein	Indels
		R- GAGGAGAACACGACGAAACAGCTGT			
RR7nr010	At3g09390	F- GGAAACTGTGGTTGTGGATCTGG	MT2a	metallothionein protein 2A	Indels
		R- ACGGGACCATCAAACTAAGGTCA			
ZIP1	At3g12750	F- GGATAGTCGTGCATTCGGTTATCAT	ZIP1	zinc transporter	Indels
		R- ATCCCTATTCCTATCCCAACCGATG			

Table 2: continued

*Tc*EST /gene	*At* ortholg	Primer	Comon Name	Description	Marker type
RR9nr034	At3g14410	F- GGGTGGTGTTGTTGGAGAAGCC	None	transporter-related	Indels
		R- GAGAGCGCTCGTGTGAGAAATCAC			
RR4nr003	At3g19820	F- TGTTCCTCTTTACAAGGTCGGCG	DWF1	cell elongation protein	Indels
		R- TCCTTGCCTCTTCTCGTACTCGAAC			
RR27nr036	At3g25940	F- GTGATGGAGAAATCTCGGGAAAGC	TFIIS	transcription factor S-II	Indels
		R- GCTCAGCATCTTCCTGCGTTTT			
RR22nr089	At3g26310	F- TCCGTTGTTGTAGTCGTCCAGGC	CYP71B35	cytochrome P450 family protein	CAPS - *Hind*III
		R- CTTCGTCCTTGATGGGTTGAATCG			
MRP10	At3g62700	F- GTTTGGAGATAAGACAGAGATCGGTGG	MRP10	glutathione-conjugate transporter 10	Indels
		R- GCACAATCATCCCATCCCGC			
RR15nr039	At4g00430	F- TGGGAGTCAAGAGAGCACCCAAC	TMP C	plasma membrane intrinsic protein	Indels
		R- TCCGAGCCAACAACAGTCCG			
RR19nr063	At4g02370	F--LC CGCTCGTGGGTTCTTACCAAC	None	expressed protein	Allele-specific
		F--GA CGCTCGTGGGTTCTTACCAAT			
		R- CCCACATTGAGGCGATTCCAA			

Table 2: continued

*Tc*EST /gene	*At* ortholg	Primer	Comon Name	Description	Marker type
RR21nr055	At4g04620	F- GTGATTGTGGAAAAAGCTGGACAG	APG8b	autophagy 8b	Indels
		R- TGGGGGTAATGTGTTCTTGACAA			
RR4nr090	At4g14710	F- GGATCAGAGACTTCCTCACCACAA	None	iron-deficiency-responsive protein	Indels
		R- GGTTGCGATCCCGAACATCA			
Tc con 257	At4g14880	F- TGGAAAGGCACTGGTGGAAA	OAS1	O-acetylserine sulfhydrylase (OAS1)	Indels
		R- TGTCGCGTCGAATAGTACAGTCG			
RR5nr073	At4g17340	LC: F- CTTCGTTACTAATGGCGAGGTACAC	DELTA	Tonoplast intrinsic protein 2;2	Allele-specific
		GA: F- CTTCGTTACTAATGGCGAGGTACAG	TIP2/TIP2;2		
		R- TGCAAAGGTCACAACAATCTCCA			
IRT2	At4g19680	F- TAAATTCGACGTAGCATTCTAAT	IRT2	iron-responsive transporter	Indels
		R- GAGATAGTCCAATGACCACAG			
RR20nr064	At4g22220	F- CCACTGAATGGGTGAAAGGCAA	ISU1	iron-sulfur cluster assembly complex	Indels
		R- TGGAGCTTCATTGGTGGGAGA			
RS13nr081	At4g25640	F- GAGTATCATTGTTCTTACTGGTCGTCTTG	None	MATE efflux family protein	Indels
		R- GGTCTCTTGGTATAATAATCGCCACC			

Table 2: continued

*Tc*EST /gene	*At* ortholg	Primer	Comon Name	Description	Marker type
RR1nr094	At4g26050	F- ACGGAATCACGGTGTTGCCG	None	leucine-rich repeat family protein	Indels
		R- TGCTTCACTGCTTCCAAACCCTG			
FER1	At5g01600	F- GTTTGAGGAGGTGAAGAAAGCCGAT	FER1	ferritin 1	Indels
		R- GCGTACATTGAATGGTACACATACG			
RR7nr085	At5g02380	F- GAGAAAATGTCTTGCTGTGGAGGAAAC	MT2b	metallothionein protein 2B	Indels
		R- CGCCGGCGAAATCAACCATT			
RR5nr078	At5g10450	F- GCTTTTGAAGAAGCCATAGCTGAG	GRF6	14-3-3 protein GF14 lambda	Indels
		R- GCAGATAAATCATCAAGCTGTGCAC			
RR3nr075	At5g21274	F- TCAGAGTTCAAGGAAGCGTTTAGCC	CAM6	calmodulin-6	CAPS - *Eco32*I
		R- CATCACGGTCCCAAGCTCCTTC			
RR25nr081	At5g23140	F- TACCTCGAGTCTGAGAATCCTTCCAAG	CLPP2	CLP protease proteolytic subunit 2	CAPS - *Tru*I
		R- GGCCTAAACAAATTGTGCTAATTGGAG			
RR1nr041	At5g25090	F- GCGCCTGGAAAATCCCTTCTTC	None	plastocyanin-like domain-containing protein	Indels
		R- GCTTCTTTTGTCACTTGCAGAACCG			
RR3nr010	At5g43780	F- CGTTGTCAACATGTCGGTTCCG	APS4	sulfate adenylyltransferase 4	Indels
		R- TCGCCTCCGATCAACCAGTTTC			

Table 2: continued

*Tc*EST /gene	*At* ortholg	Primer	Comon Name	Description	Marker type
RR10nr088	At5g46710	F- TTGCTTAGACTGCACCAATGGC	None	zinc-binding family protein	Indels
		R- GGTTCCATTGGTGTATCCTTCACTC			
RS5nr034	At5g53330	F- GGAATCAAGGTGACTCTGAAGCCT	None	expressed protein	Indels
		R- CCAATTCAAGAAGAGCCATGGAG			
RS26nr082	At5g67330	F- AATCGCCGGGTACCGGAAAG	NRAMP4	NRAMP metal ion transporter 4	CAPS - *Mse*I
		R- TTACAACTCCAGCCCAGAGAGGAATC			
RS16nr083	None	F- TTTCTTCTCAGCAAAACTCAGTCTCCG	None	None	Indels
		R- AAGAAAACATGGTGACATGGACATGC			
RR6nr032	None	F- GGATGATAAGATCGAACCTGAAGAGGC	None	None	Indels
		R- TGGTTCCATAACCACAAGAGAACTTGG			

Table 2: continued

Each EST sequence was compared to the *At*-genome using BLASTN (Altschul *et al.,* 1990), to identify the putative locations of introns and Cleaved Amplified Polymorphic Sequence (CAPS) markers were designed if the polymorphism removed or created a restriction site. Large (>15 bp) insertions or deletions (indels) were scored on 2-3% agarose or MetaPhor agarose gels (Cambrex Corporation), whereas acrylamide gels were used to separate small indels (<15 bp) (Table 3).

AK Block	~Interval	*Al* LG	*Al* order	*Al* orient.	*At* LG	*At* order	*At* orient.	*Tc* LG	*Tc* order	*Tc* orient.	Brassica block
A	At1g01560-At1g19330	1	1	+	1	1	+	2	3	+	C1A
B	At1g19850-At1g36240	1	2	+	1	2	+	2	4	+	C1B
C	At1g43590-At1g56145	1	3	+	1	3	+	2	5	+	C1C
D	At1g63770-At1g56520	2	4	+	1	4	-	1	1	(+)	C1D
E	At1g65040-At1g80420	2	5	+	1	5	+	1	2	-	C1E
F	At3g01040-At3g25520	3	6	+	3	11	+	7	24	+	C3A
G	At2g05170-At2g07733	3	7	+	2	7	+	5	18	(+)	C2A
H	At2g15670-At2g21140	3	8	+	2	8	+	5	19	(+)	C2A
I	At2g21160-At2g28910	4	9	+	2	9	+	5	20	(+)	C2B
J	At2g31040-At2g47730	4	10	+	2	10	+	5	21	+	C2C
K	At2g01250-At2g03750	5	11	+	2	6	+	5	17	(+)	C2A
L	At3g25855-At3g29772	5	12	+	3	12	+	4	14	+	C3B
M	At3g43740-At3g49970	5	13	+	3	13	+	4	15	+	C3C
N	At3g50950-At3g62790	5	14	+	3	14	+	4	16	+	C3D
O	At4g00030-At4g04955	6	15	+	4	15	+	3	6	+	C4A
P	At4g12070-At4g08690	6	16	+	4	16	-	3	7	(-)	C4A
Q	At5g28897-At5g22800	6	17	+	5	20	-	4	13	+	C5B
R	At5g22030-At5g01240	6	18	+	5	19	-	3	8	+	C5A
S	At5g41900-At5g32621	7	19	+	5	21	-	4	12	(-)	C5C
T	At4g12750-At4g16143	7	20	+	4	17	+	6	23	-	C4B
U	At4g16250-At4g38770	7	21	+	4	18	+	6	22	-	C4B
V	At5g48520-At5g42970	8	22	+	5	22	-	4	11	+	C5D
W	At5g49430-At5g60390	8	23	+	5	23	+	4	10	-	C5E
X	At5g60550-At5g67385	8	24	+	5	24	+	4	9	-	C5F

Table 3: 'Ancestral Karyotype' block summary, modified from Schranz *et al.* (2006). Hypothetical block orientation is indicated in brackets. *Al* = *Arabidopsis lyrata*; *Al* order = Order of blocks along *A. lyrata* LGs; *Al* orient. = relative orientation of blocks along *A. lyrata* LGs. *At* = *Arabidopsis thaliana*; *Tc* = *Thlaspi caerulescens*,

Some of these co-dominant markers correspond to genes known to be involved in metal homeostasis, such as *TcZNT1* (Pence *et al.*, 2000), *TcZNT2* (Assunção *et al.*, 2001), *TcZIP1* (Grotz *et al.*, 1998; van de

Mortel *et al.*. 2006), *TcZIP5* (van de Mortel *et al.*, 2006), *TcFER1* (Petit *et al.*, 2001), *TcMT2a* (Zhou and Goldsbrough, 1995; van de Mortel *et al.*, 2006), *TcMT2b* (Zhou and Goldsbrough, 1995; van de Mortel *et al.*, 2006), *TcNRAMP3* (Thomine *et al*. 2000; van de Mortel *et al.*, 2006), *TcNRAMP4* (Thomine *et al*. 2000), *TcZTP1* (Assunção *et al.*, 2001), *TcIRT2* (Vert *et al.*, 2001).

Mapping analysis

Marker segregation data were obtained by analyzing the entire F_3 population for AFLP markers using 16 selected *E/M* and *P/M* PCs together with the CAPS and Indel markers. Linkage analysis was performed with the computer software package JOINMAP 3.0 (Stam 1993; Stam and van Ooijen 1996). Kosambi's mapping function (Kosambi 1944) was used to convert recombination frequencies into map distances (cM). The pair-wise analysis obtained from JOINMAP was used to assign markers to linkage groups (LGs) with Log of Odds (LOD) score ≥4. When the combined map was compared with each of the parental maps, in general a similar marker order was found (data not shown).

QTL mapping

Potential QTLs for each trait were identified using the MAPQTL 5.0 package (van Ooijen, 2004). Kruskal-Wallis and interval mapping (IM) analyses were initially performed to find regions with potential QTL effects. Co-dominantly, otherwise dominantly, scored markers were then used in various combinations as co-factors in multiple QTL models in

Multiple QTL Model analysis (MQM analysis also performed with MAPQTL). Using MAPQTL software, permutation tests calculated LOD score thresholds for each trait which was applied to declare the presence of a QTL. This corresponds to a general genome-wide significance of $P < 0.05$ for normally distributed data, as was determined by extensive simulation experiments (van Ooijen, 1999). The QTL graphs were prepared with MAPCHART (Voorrips, 2002).

Results

Linkage map construction

144 F_3 individuals were genotyped for 95% of all the dominant and co-dominant markers, which were used for the map construction. A total of 170 AFLP markers (103 GA-specific and 67 LC-specific) were scored using the 16 selected PCs. In addition, 44 co-dominant markers (nine CAPS and 35 indels) were also scored in the F_3 mapping population. An integrated map was constructed using both dominant and co-dominant markers (Figure 1). All the markers were assigned to seven linkage groups (LG1-LG7), using a LOD score of 4.0. The lengths of the linkage groups, corresponding to the number of haploid chromosomes, varied from 48 to 90 cM. The seven LGs were ordered according to common markers with the genetic map published by Deniau *et al.* (2006). The total map length was 521 cM with an average distance between two adjacent markers of 2.45 cM and a largest distance of 19 cM (LG6). The marker density per cM was not significantly different between LGs. In each of the LGs, there was a cluster of linked markers, probablyrepresenting the centromeric regions (Figure 1). Most of the loci showed genotypic ratios as expected for a F_3 segregating population (0.375 : 0.25 : 0.375). After correction according to the Bonferroni-Holm

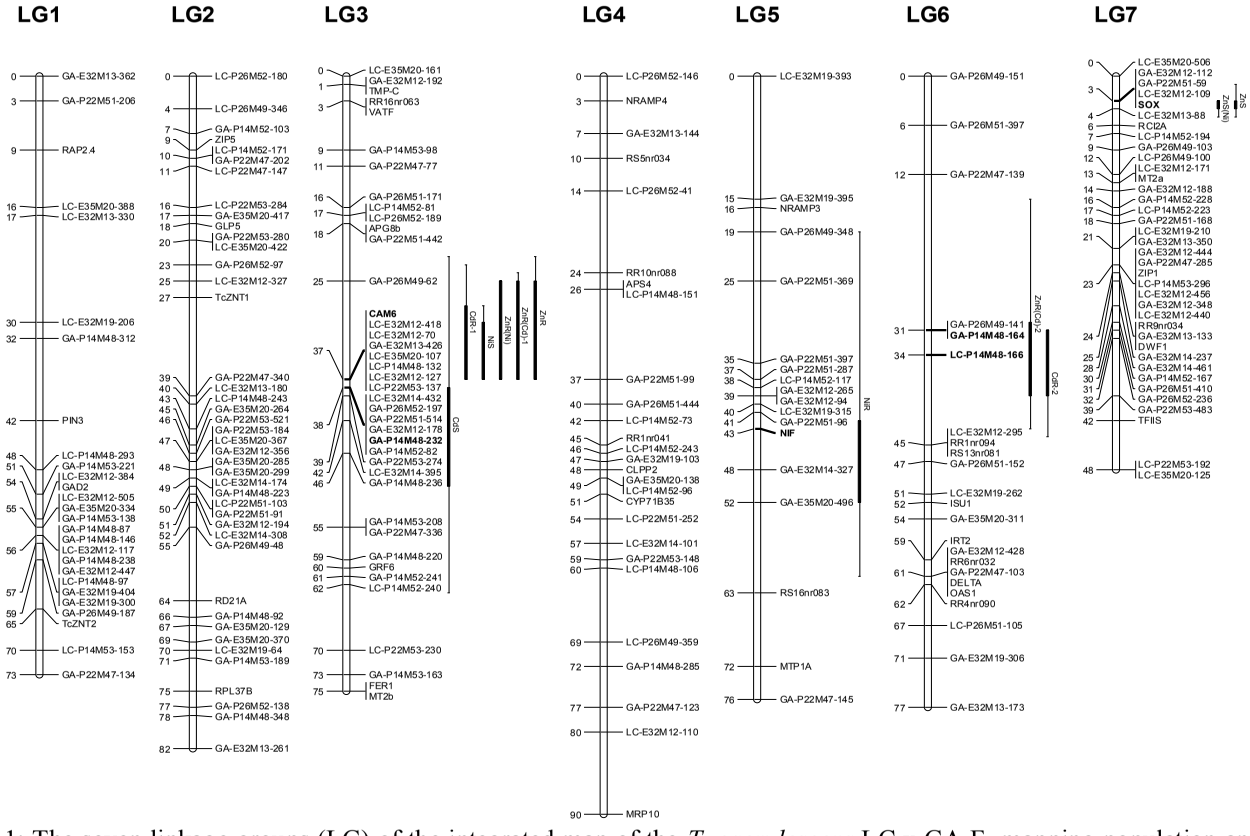


Figure 1: The seven linkage groups (LG) of the integrated map of the *T. caerulescens* LC x GA F_3 mapping population and the location of all significant QTLs for all the traits analyzed supported by their 1-LOD (*bars*) and 2-LOD (*lines*) support intervals. The genetic distances are given in cM. The MML are indicated in *bold*.

Comparative analysis of the *Thlaspi caerulescens* map with a unified comparative genomic framework of the Brassicaceae

Recent studies suggest that comparison with an 'ancestral karyotype' (AK) of $n = 8$ would considerably expedite genomic comparison (Koch and Kiefer, 2005; Lysak *et al.*, 2006). Schranz *et al.* (2006) proposed a set of 24 genomic blocks (A-X) within the concept of the $n = 8$ AK shared by *A. lyrata* and *C. rubella* (Lysak *et al.*, 2006) in order to create a unified comparative genomics system across the Brassicaceae. The transition from eight chromosomes of the AK map to seven chromosomes in the *Thlaspi caerulescens* map can be explained by nine main chromosomal rearrangements (four translocations, two fusions and three inversions). These consist of an inversion of block V (AK8) followed by a translocation of block S (AK7) to the inverted AK8. Block Q (AK6) was translocated and inverted to block S. A fusion occurred between the blocks O-P (AK6) and block R (AK6). The blocks G-H (AK3) were translocated to AK4 and block K (AK5) to block G. The rest of AK5 (blocks L-M-N) was fused to block Q. In other words, the *Thlaspi caerulescens* map is composed of the inverted AK2 (LG1), AK1 (LG2), of blocks O-P and R (LG3), of the inverted AK8 fused with block S and the inverted block G and the blocks L-M-N (LG4). LG5 consists of blocks K and G-H and AK4, LG6 is blocks T-U and LG7 is block F (Table 3, Figure 2). sequential method (Rice 1989), the ratio of 3 markers was significantly distorted (Figure 1). These markers were closely linked on LG1 between 56 and 57 cM.

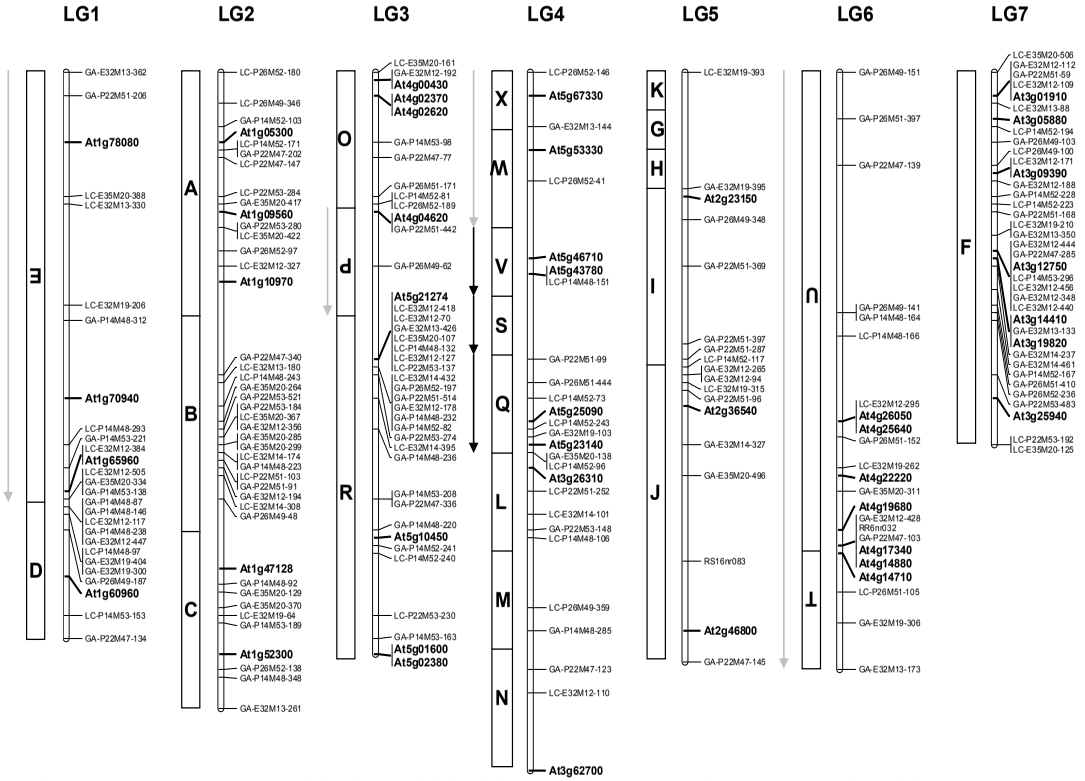


Figure 2: 'Ancestral Karyotype' (AK; n=8) blocks and block boundaries mapped onto the deduced karyotype (n = 7) of *Thlaspi caerulescens*. The genomes blocks defined by their position in the AK (Table 3) are reorganized to show the evolution of the reduced karyotype of *T. caerulescens* (n = 7). Blocks that have been inverted relative to the AK are represented by black downward-pointing arrows (blocks V, S and Q). Blocks which are in opposite orientation relative to the AK, but not inverted, are represented by gray downward-pointing arrows and by the block letter being upside-down (blocks E, P, T, U, W and X).

Characterization of Cd, Zn and Ni accumulation in LC, GA and the F_3 progeny

The trait abbreviations have been named according to the metal treatment followed by the tissue analyzed (for example, Cd accumulation in root = CdR). Zinc accumulation in root and in shoot, respectively, were also measured in Cd and Ni treatments. Therefore, we added in brackets the metal treatment (for example, Zn accumulation in shoot in Cd treatment = ZnS(Cd)). For the Cd treatment, we performed the same experiment as previously described (Deniau *et al.*, 2006) with the exception that Zn accumulation in shoot and in root, as described by Deniau *et al.* (2006), are now called ZnS(Cd) respestively ZnR(Cd). The CdS, CdR, ZnS(Cd) and ZnR(Cd) parental distributions were already described by Deniau *et al.* (2006). In the Ni treatment (Table 4), the ranges of NiS and NiR for the GA accession (0.74-13.21 and 0.53-9.49 $\mu mol.g^{-1}$ DW, respectively, n = 52) were slightly overlapping with those from the LC accession (0.74-9.55 and 0.53-6.36 $\mu mol.g^{-1}$ DW, respectively, n = 46). This was also the case for ZnR(Ni) (0.85-17.26 $\mu mol.g^{-1}$ DW for GA and 0.51-8.93 $\mu mol.g^{-1}$ DW for LC) and we found a broad overlap for ZnS(Ni) [GA (2.57-14.86 $\mu mol.g^{-1}$ DW), LC (0.45-39.55 $\mu mol.g^{-1}$ DW)]. On average, NiS, NiR, ZnS(Ni) and ZnR(Ni) were 2.1 -, 3.5-, 1.4- and 6.1 -fold higher in GA than in LC, respectively (Table 4). In the Zn treatment, ZnS ranged from 29.6 to 327 $\mu mol.g^{-1}$ DW for GA and 7.9 to 95 $\mu mol.g^{-1}$ DW for LC. We observed only a little overlap for ZnR [GA (4.9-246 $\mu mol.g^{-1}$ DW), LC (1.9-69.2 $\mu mol.g^{-1}$ DW)]. ZnS and ZnR were 3.8-fold higher and 8.2-fold higher in GA than in LC, respectively (Table 4).

Trait	Population / accession	Median (in µmol.g^{-1} DW)	95% confidence intervals (in µmol.g^{-1} DW)
CdS	LC	0.51	0.44 - 0.60
	GA	3.41	2.94 - 3.95
	F_3	1.36	1.05 - 1.76
CdR	LC	0.62	0.55 - 0.71
	GA	7.52	6.98 - 8.09
	F_3	2.34	2.08 - 2.63
ZnS(Cd)	LC	4.24	3.87 - 4.65
	GA	2.72	2.51 - 2.94
	F_3	3.31	2.90 - 3.78
ZnR(Cd)	LC	1.18	1.13 - 1.23
	GA	2.71	2.08 - 3.52
	F_3	0.91	0.84 - 0.99
NiS	LC	2.14	1.89 - 2.42
	GA	4.59	4.01 - 5.26
	F_3	3.21	2.80 - 3.67
NiR	LC	1.05	0.92 - 1.19
	GA	3.68	3.38 - 4.00
	F_3	2.04	1.72 - 2.42
ZnS(Ni)	LC	4.29	3.59 - 5.14
	GA	5.83	5.20 - 6.53
	F_3	5.05	4.44 - 5.74
ZnR(Ni)	LC	0.98	0.86 - 1.12
	GA	6.02	5.12 - 7.06
	F_3	2.57	2.04 - 3.24
ZnS	LC	31.38	26.97 - 36.51
	GA	113.96	95.16 - 136.47
	F_3	60.61	51.18 - 71.78
ZnR	LC	8.49	7.02 - 10.27
	GA	69.52	53.81 - 89.90
	F_3	24.84	19.67 - 31.35

Table 4: Median trait levels and 95% confidence intervals for the La Calamine (LC) and Ganges (GA) accessions and the F_3 population. CdS (Cd shoot accumulation), CdR (Cd root accumulation), ZnS(Cd) (Zn accumulation in shoot in Cd treatment), ZnR(Cd) (Zn accumulation in root in Cd treatment), NiS (Ni shoot accumulation), NiR (Ni root accumulation), ZnS(Ni) (Zn accumulation in shoot in Ni treatment), ZnR(Ni) (Zn accumulation in root in Ni treatment), ZnS (Zn shoot accumulation), ZnR (Zn root accumulation),

QTL mapping

The genetic map containing 214 dominant and co-dominant markers was used for QTL analysis. All 10 traits were analyzed and for all traits except for ZnS (Cd) at least one QTL was detected. The 11 QTLs with significant effects were distributed over four LGs (Figure 1). Individual

	CdS	CdR	ZnS(Cd)	ZnR(Cd)	NiS	NiR	ZnS(Ni)	ZnR(Ni)	ZnS	ZnR
CdS	-	0.29***	0.59***	-0.24 *ns*	0.22**	0.02 *ns*	0.16 *ns*	0.04 *ns*	0.09 *ns*	0.08 *ns*
CdR	-	-	0.18**	0.40***	0.35***	0.11 *ns*	-0.03 *ns*	0.41***	0.18*	0.43***
ZnS(Cd)	-	-	-	-0.08 *ns*	0.12 *ns*	-0.02 *ns*	0.35***	0.01*ns*	0.12 *ns*	-0.01 *ns*
ZnR(Cd)	-	-	-	-	0.27**	0.09 *ns*	0.23**	0.37***	0.33***	0.37***
NiS	-	-	-	-	-	0.29**	0.66***	0.48***	0.53***	0.45***
NiR	-	-	-	-	-	-	0.06 *ns*	0.55***	0.17**	0.39***
ZnS(Ni)	-	-	-	-	-	-	-	0.18*	0.55***	0.26**
ZnR(Ni)	-	-	-	-	-	-	-	-	0.52***	0.79***
ZnS	-	-	-	-	-	-	-	-	-	0.65***
ZnR	-	-	-	-	-	-	-	-	-	-

Table 5: Pearson correlations of the log-transformed trait data. CdS (Cd shoot accumulation), CdR (Cd root accumulation), ZnS(Cd) (Zn accumulation in shoot in Cd treatment), ZnR(Cd) (Zn accumulation in root in Cd treatment), NiS (Ni shoot accumulation), NiR (Ni root accumulation), ZnS(Ni) (Zn accumulation in shoot in Ni treatment), ZnR(Ni) (Zn accumulation in root in Ni treatment), ZnS (Zn shoot accumulation), ZnR (Zn root accumulation). Significance of the correlation: *ns* not significant; * $P < 0.1$; ** $P < 0.05$; *** $P < 0.001$

QTLs accounted for 8.5-40% of the phenotypic variation in this population (Table 6). Six QTLs located on LG3, LG5, LG6 and LG7 were identified by Interval Mapping (IM) and subsequent Multiple QTL Model (MQM) analysis. The QTLs were named after the trait abbreviation. Only for Zn accumulation in root while exposed to Cd and for Cd accumulation in root we found two QTL (ZnR(Cd)-1 on LG3, ZnR(Cd)-2 on LG6; CdR-1 on LG3 and CdR-2 on LG6), explaining respectively 8.9, 9.7, 21.4 and 14.1% of the phenotypic variance respectively, with a multiple QTL PVE (Percentage Variance Explained)

of 21.3% for Zn and 29.7% for Cd accumulation (Table 5). Interestingly, CdR-1 and ZnR(Cd)-1 are partly co-locating and share the same most closely associated Molecular Marker Loci (MML). Although CdR-2 and ZnR(Cd)-2 do not have the same MML they are also partly co-locating and their 1-LOD support intervals are spanning the same genomic region. Due to the lack of co-dominant markers with sufficient linkage to CdR-2 and ZnR(Cd)-2, we were unable to establish any epistatic interactions between QTLs. CdR-1 and ZnR(Cd)-1 are part of a group of five (co-locating) QTLs linked to the same MML. This cluster also contains NiS, ZnR and ZnR(Ni) explaining 40, 26 and 31% of the trait variance respectively. CdS is also located on LG3 (8.5%). Two more QTLs (ZnS and ZnS(Ni)) sharing the same MML were detected on LG7 and accounted for 14.4 and 25.1% of the phenotypic variance, respectively. Only one QTL, NiR, was detected on LG5 and explained 14.1% of the variance in Ni accumulation in root. For all the QTLs, except NiR, the trait-enhancing alleles originate from GA.

Metal treatment	Trait	GW LOD threshold	QTL	LG	MML	LOD score	% PVE	Multiple QTL PVE (%)	Trait-enhancing allele
Cd	CdS	2.4	CdS	3	GA-P14M48-232	2.7	8.5		GA
	CdR	2.3	CdR-1	3	At5g21274	10.2	21.4	29.7	GA
			CdR-2	6	LC-P14M48-164	7.1	14.1		GA
	ZnR(Cd)	2.3	ZnR(Cd)-1	3	At5g21274	3.6	8.9	21.3	GA
			ZnR(Cd)-2	6	LC-P14M48-166	3.9	9.7		GA
Ni	NiS	2.3	NiS	3	At5g21274	9.9	40.0		GA
	NiR	2.6	NiR	5	At2g36540	2.9	14.1		LC
	ZnS(Ni)	2.5	ZnS(Ni)	7	At3g01910	5.7	25.1		GA
	ZnR(Ni)	2.4	ZnR(Ni)	3	At5g21274	7.4	31.0		GA
Zn	ZnS	2.4	ZnS	7	At3g01910	3.0	14.4		GA
	ZnR	2.5	ZnR	3	At5g21274	5.8	26.0		GA

Table 6: Overview of QTLs detected for all the analyzed traits. QTLs are named according to their trait abbreviations followed by a number distinguishing QTLs affecting the same trait. Trait-enhancing allele = the allele that has a positive effect on the trait value. GW LOD threshold indicates genome wide log-likelihood threshold determined by MapQTL5 permutation test (n = 1.000). Multiple QTL PVE indicates total phenotypic variance explained by all detected QTLs. % PVE = percentage of explained phenotypic variance.

Discussion

Linkage map reliability and synteny to other Brassicaceae

The current map of *Thlaspi caerulescens* compares very well to the genetic maps previously published (Assunção *et al.*, 2006; Deniau *et al.*, 2006). Both the total map length and the sizes of the corresponding LGs were similar. A comparison based on common AFLP and *Arabidopsis thaliana*-anchored markers showed the same map positions and no discrepancy in marker order in both LC x GA maps. While overall the AFLP markers are evenly distributed, some cluster in distinct regions, corresponding with the centromeric regions of the *T. caerulescens* chromosomes. Also, the co-dominant markers based on evenly spread Arabidopsis genes are well distributed over the different LGs, except for the top of LG6, where co-dominant markers are lacking.

The number of orthologous markers enables us to make further assumptions on the whole genome synteny and to establish chromosomal rearrangements as was done for another Zn and Cd hyperaccumulator *Arabidopsis halleri* (Willems *et al*., 2007) and the related *A. lyrata* (Kuittinen *et al.*, 2004; Koch and Kiefer, 2005; Yogeeswaran *et al.*, 2005). These comparative mapping studies have shown that both $n = 8$ genomes are largely collinear with the reduced $n = 5$ genome of *A. thaliana*. This is also observed in the *T. caerulescens* comparative map. In addition, Lysak *et al.* (2006) also examined other karyotypes with reduced chromosome numbers ($n = 6$ and 7) of two taxa from the tribe Camelineae (*Neslia*, *Turritis*) and one taxon from Descurainieae (*Hornungia*).

The genome sizes reported for *A. thaliana* (0.16 pg) and its close relative *A. halleri* (~0.26 pg) (Johnstone *et al*., 2005) and *T. caerulescens* (~0.30 pg) (Peer *et al*., 2006) indicate that one or more deletion events might have accompanied the transition from eight to seven chromosomes (*T. caerulescens*) and from eight to five chromosomes, characterizing the genome of *A. thaliana*. Ideally, the comparative analyses of the *Ah x Alp* (Willems *et al.,* 2007), the *A. lyrata* (Kuittinen *et al*., 2004; Yogeeswaran *et al*., 2005) and the *T. caerulescens* maps with the *A. thaliana* genome should take these events into account.

Overall, the *T. caerulescens* comparative map and the results obtained by Lysak et al. (2006) revealed that all species analyzed shared conserved segments that can be related to the AK (Table 3). This comparative genomic study will facilitate to further unravel the identified QTLs, by identification of the regions of the AK/Arabidopsis genome the *T. caerulescens* QTL regions correspond to and to use this information to either fine-map the regions or search directly for possible candidate genes based on a presumed function in metal homeostasis. Also, in the field of heavy metal hyperaccumulation and tolerance, this comparative mapping will be the key for the transfer of knowledge from model species to other Brassicaceae with economic and/or agronomic applications.

Segregation for Cd, Ni, Zn, Zn(Cd) and Zn(Ni) hyperaccumulation traits

In addition to the results of Deniau *et al.* (2006) on the accumulation of Cd and Zn, we observed that GA accumulates more Ni in roots and

shoots and also more Zn, both in the presence or absence of Ni. For Ni, these results compare well with the observations made by Assunção *et al.* (2003c). The slight discrepancy between the Ni accumulation data might be due to the time of exposure (3 weeks in Assunção *et al.*, 2 weeks in ours) or the concentration of Zn in to the nutrient solution (1 µM in Assunção *et al.*, 2 µM in ours).

The frequency distributions of the Cd, Zn(Cd), Ni, Zn(Ni) or Zn concentrations in shoot and in root in the F_3 population are not bimodal suggesting either that all these traits are quantitatively inherited, or that the heritabilities are low. Similar continuous distributions were also found fr Cd and Zn accumulation in other segregating populations (Assunção *et al*., 2003b; Zha *et al*., 2004; Assunção *et al*., 2006; Deniau *et al.*, 2006).

For all traits we observed that the alleles enhancing accumulation were co-dominant or recessive, which is in accordance with previous analyses of *T. caerulescens* inter-accession crosses reported so far, in which F_1 plants showed accumulation rates that were either intermediate between the parents, or directed towards low accumulation (Frérot *et al*., 2003, 2005; Assunção *et al*., 2006).

QTL analysis

The QTL analysis often identified only one locus contributing to a trait, although phenotypic segregation suggested multiple loci. This suggests either that the heritability is not very high, or that there are so many loci contributing to the traits that each locus has a relatively small effect, and is thus harder to detect. The (multiple) QTL PVE ranged from 8.5 to

40% (Table 6). We confirmed the detection of a QTL for CdS on LG3 explaining no more than 8.5% of the phenotypic variance (Figure 1, Table 5). This QTL is located at the same position as the one detected in the F_2 population (Deniau *et a*l., 2006). However, the low PVE is rather surprising compared to the one found in the F2 (22.8%). This is possibly due to the smaller size of the F_2 population (n = 129) and/or a large number of minor QTLs that could be detected in a larger population. The so-called "Beavis effect" predicts that in larger populations more QTLs of low effect are detected than in smaller populations (Beavis, 1994; Kearsey and Farquhar, 1998; Xu, 2003). Also, an inflation of the estimates of the detected QTLs occurs in a population of intermediate size, such as the one used in this study (n = 145). This could explain the low genetic variance explained by this QTL.

The NiR QTL on LG5 explained 14.1% of the phenotypic variance. The trait-enhancing allele of this QTL is the only one originating from LC. Using the synteny between *T. caerulescens* and *A. thaliana*, the putative candidate gene will be in the vicinity of NIF (At2g36540; Table 2). In Arabidopsis this region will contain the *IREG1* gene, which is 84.6% similar on nucleotide level to *AtIREG2*. Van de Mortel (2006) has shown that this gene is overexpressed in *T. caerulescens* vs *A. thaliana*. Recently this gene was described to encode a tonoplast transport protein involved in iron-dependent nickel detoxification in Arabidopsis roots (Schaaf *et al*., 2006). These authors propose a role for *AtIREG2* in vacuolar loading of nickel under iron deficiency and their results identify *AtIREG2* as a yet unrecognized component in metal homeostasis required to circumvent toxicity of nickel in iron-deficient plants. In

addition, *AtIREG1* (At2g38460) and *AtIREG2* share 84.6% in sequence similarity which suggest a similar role in Ni detoxification. Unfortunately, the role of these two genes in hyperaccumulators has not been described yet and further studies are necessary to determine if/how these proteins function to hyperaccumulate Ni *in planta.*

Interestingly, we detected three cases of QTL co-location (Figure 1): one for Zn accumulation in shoots, located on LG7 (ZnS and ZnS(Ni)), one for Zn and Cd accumulation in roots, located on LG6 (ZnR(Cd)-2/CdR-2) and one of Zn and Cd accumulation in roots and Ni accumulation in shoots, located on LG3 (NiS, CdR-1, ZnR(Cd)-1, ZnR and ZnR(Ni)). Although we do not yet have direct evidence that these QTLs are actually caused by variation in the same gene, it is tempting to speculate about possible reasons. The co-location of Zn and Zn(Ni) shoot accumulation on LG7 suggests that Ni has little or no effect on Zn accumulation in *Thlaspi caerulescens*, a such as experimentally demonstrated by Assunção *et al*. (2001, 2008). On LG6, the co-location of CdR-2 and ZnR(Cd)-2 was also mapped on the same LG in the study of Deniau *et al.* (2006). However, in the previous map, these QTLs were located to the top of LG6. We believe they may map at the same position since the LG6 of the present genetic map is covering more genome (61 and 77 cM, respectively). Thus, the co-location of these QTL was detected in two independent experiments, with both trait-enhancing alleles originating from GA in both cases. This may be taken to confirm the existence of a Zn/Cd transporter with a relative high affinity for Cd and low affinity for Zn, such as proposed to be present in the Cd-accumulation ecotype GA (Zha *et al*., 2004). The lack of Arabidopsis-

anchored markers at the QTL's position does not enable us to speculate on candidate genes.

The cluster of five QTLs on LG3 consists of three QTLs related to Zn accumulation in root (ZnR, ZnR(Cd)-1 and ZnR(Ni)), one for Cd accumulation in root (CdR-1) and also one for Ni accumulation in shoot (NiS). These five QTLs all co-segregate with the CAM6 marker (Table 2) and their trait-enhancing alleles are all originating from GA, suggesting that indeed the same gene is involved. In the previous study of Deniau *et al.* (2006), CdR-1 was also found to co-segregate with CAM6. According to the QTL support intervals and their local synteny, the gene(s) underlying these QTLs is (are) located at the end of block R (At5g21274-At5g22030) or in block P (At4g12070-At4g08690). One of the genes in this region with a GO annotation suggesting an involvement in metal transport is *AtAHA12* (At4g11730), which encodes a P-type ATPase metal transporter involved in cation transport. Assunção *et al.* (2006) previously mapped QTLs for Zn accumulation in root in a *T. caerulescens* F_3 inter-accession cross using a common parent (LC) with the cross used in this study. There were two significant QTLs detected for this trait. Remarkably, the QTL, named LE3, co-segregated with a co-dominant marker (indel47/48), representing At5g21274 (calmodulin 6 / CAM6). In our study, the Zn accumulation in root-related traits (ZnR(Cd)-1, ZnR(Ni) and ZnR) on LG3 are also co-segregating with CAM6. Interestingly, the trait-enhancing allele of LE3 (Assunção *et al.*, 2006) and the Zn accumulation in root related traits on LG3 (this study) are originating from the non-common parents of the crosses, LE and GA respectively. The gene underlying these QTLs might be the same.

It is remarkable that we did not find common QTLs for Zn and Ni accumulation either in roots or in shoots, since Assunção *et al*. (2001, 2008) provided evidence of a common accumulation system for these metals, which is present in GA, but not or barely in LC. In addition, Richau and Schat (2009) established a high degree of genetic correlation between Zn and Ni accumulation in the leaves in a LC x MP inter-accession cross. However, MP (Monte Prinzera, Italy) is a much better Ni hyperaccumulator than GA. In the present study we did find a common QTL for Ni, Zn and Cd accumulation, on LG3, but this QTL controlled Zn and Cd accumulation in root, but Ni accumulation in shoot. As yet we don't have any explanation for this.

Orthologs of genes that have previously been described to be involved in metal (hyper)accumulation, such as *TcZNT2* (LG1) (Assunção *et al.*, 2001), *TcZNT1* (LG2) (Pence *et al.*, 2000; Assunção *et al.*, 2001), *TcZIP5* (LG2) (van de Mortel *et al.*, 2006), *TcFER1* (LG3) (Petit *et al.*, 2001), *TcMT2b* (LG3) (Zhou and Goldsbrough, 1995; van Hoof *et al.*, 2001; van de Mortel *et al.*, 2006), *TcNRAMP4* (LG4) (Thomine *et al*. 2000), *TcNRAMP3* (LG5) (Thomine *et al*. 2000; van de Mortel *et al.*, 2006), *TcZTP1* (LG5) (Assunção *et al.*, 2001; Dräger *et al.*, 2004), *TcIRT2* (LG6) (Vert *et al.*, 2001), *TcZIP1* (LG7) (Grotz *et al.*, 1998; van de Mortel *et al.*. 2006), *TcMT2a* (LG7) (Zhou and Goldsbrough, 1995; van Hoof *et al.*, 2001; van de Mortel *et al.*, 2006) were mapped on the LC x GA *T. caerulescens* genetic map. Their location, however, did not coincide with any of the detected QTLs and therefore there does not seem to be substantial genetic variation for these genes in this population that can be associated with differences in Zn, Cd or Ni accumulation.

Recently, Filatov *et al.* (2006) conducted a comparison of gene expression in F_3 segregating families derived from an inter-specific cross between *A. halleri* and *A. lyrata* ssp. *petrea* to identify genes and genomic regions involved in Zn hyperaccumulation. Among other genes, they found a cation transporter, ZIP6 (At2g30080), which is consistently highly expressed in leaves and roots of all accumulating F_3 genotypes. This gene was also identified to be highly expressed in *A. halleri* shoots when comparing it to *A. thaliana* (Becher *et al.*, 2004). There is currently no information on its metal specificity (Guerinot, 2000); however, Filatov *et al.* suggest ZIP6 as a strong candidate for involvement in Zn hyperaccumulation. They found two QTL on chromosome 3 and 7 of *A. halleri*. These QTL were confirmed in a recent study (Filatov *et al.*, 2007). Using the synteny between *A. halleri* and *T. caerulescens*, ZIP6 is located in the vicinity of our NiR QTL (LG5). It is also close to the root Zn accumulation QTL detected on chromosome 5 in the study of Assunção *et al*. (2006). However, further fine-mapping or gene silencing experiments will be necessary to substantiate the candidature of this gene. Finally, based on the chromosomal synteny, the gene encoding the stelar heavy metal transporting ATPase, HMA4, seems to be outside all of the QTL regions found in the present study. The strongly enhanced expression of this gene in *A. halleri*, compared with *A. thaliana*, due to tandem triplication and enhanced cis-activation in the former species, has been shown to be essential for its Zn hyperaccumulation and Cd and Zn hypertolerance phenotypes (Hanikenne *et al*., 2008). However, its expression in *T. caerulescens*, which is comparably enhanced in comparison with *A. thaliana* (van de Mortel *et al*., 2006), may be rather

similar in different accessions of this species (Xing *et al*., 2008). This suggests that HMA4, although it seems to be responsible for major QTL for Cd and Zn in the inter-specific *A. halleri* x *A. lyrata* cross (Willems *et al*., 2007; Courbot *et al*., 2007), is probably not responsible for QTL in intra-specific *T. caerulescens* crosses. Similarly, many of the metal homeostasis genes that we used as co-dominant markers might be more or less similarly expressed in different *T. caerulescens* accessions (Assunção *et al*., 2001), albeit much stronger then in *A. thaliana* (van de Mortel *et al*., 2006). If so, then their cis-regulatory domains are unlikely to be responsible for QTL in intra-specific crosses between hyperaccumulator accessions.

In summary, our study proves that AFLP markers are adequate for map construction in *Thlaspi caerulescens*, which is in agreement with previous studies (Assunção *et al.*, 2006; Deniau *et al.*, 2006; for a review Meudt and Clarke, 2007). The QTL mapping confirmed some of the QTLs previously described and provides further opportunities for future fine mapping. To pinpoint candidate genes, the establishment of the chromosal synteny and co-linearity with *A. thaliana* will be very helpful.

References

Alonso-Blanco C, Koornneef M (2000) Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. Trends in Plant Science 5, 22–29

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. Journal of Molecular Biology 215, 403-410

Assunção AGL, Bleeker P, ten BookumWM, Vooijs R, Schat H (2008) Intraspecific variation of metal preference patterns for hyperaccumulation in *Thlaspi caerulescens*: evidence from binary metal exposures. Plant and Soil 303, 289-299

Assunção AGL, Da Costa Martins P, De Folter S, Vooijs R, Schat H, Aarts MGM (2001) Elevated expression of metal transporter genes in three accessions of the metal hyperaccumulator *Thlaspi caerulescens*. Plant, Cell and Environment 24, 217–226

Assunção AGL, Schat H, Aarts MGM (2003a) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. New Phytologist 159, 351–360

Assunção AGL, Ten Bookum WM, Nelissen HJM, Voijs R, Schat H, Ernst WHO (2003b) A co-segregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator. New Phytologist 159, 383–390

Assunção AGL, Ten Bookum WM, Nelissen HJM, Voijs R, Schat H, Ernst WHO (2003c) Differential metal-specific tolerance and

accumulation patters among *Thlaspi caerulescens* populations originating from different soil types. New Phytologist 159, 411-419

Assunção AGL, Pieper B, Vromans J, Lindhout P, Aarts MGM, Schat H (2006) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. New Phytologist 170, 21–32

Baker AJM, Reeves RD, Hajar ASM (1994) Heavy metal accumulation and tolerance in British populations of the metallophyte *Thlaspi caerulescens* J&C Presl (Brassicaceae). New Phytologist 127, 61–68

Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. Pp. 250-266 in Proceedings of the 49th Annual Corn and Sorghum Industry Research Conference, edited by A.S. Trade, Washington D.C.

Becher M, Talke IN, Krall L, Kramer U (2004) Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. Plant Journal 37, 251-268

Courbot M, Willems G, Motte P, Arvidsson S, Roosens N, Saumitou-Laprade P, Verbruggen N (2007) A major quantitative trait locus for cadmium tolerance in *Arabidopsis halleri* colocalizes with HMA4, a gene encoding a heavy metal ATPase. Plant Physiology 144, 1052-1065

Deniau AX, Pieper B, Ten Bookum WM, Lindhout P, Aarts MGM, Schat H (2006) QTL analysis of cadmium and zinc accumulation in

the heavy metal hyperaccumulator *Thlaspi caerulescens*. Theoretical and Applied Genetics 113, 907–920

Dräger DB, Desbrosses-Fonrouge AG, Krach C, Chardonnens AN, Meyer RC, Saumitou-Laprade P, Kramer U (2004) Two genes encoding *Arabidopsis halleri* MTP1 metal transport proteins co-segregate with zinc tolerance and account for high MTP1 transcript levels. Plant Journal 39, 425-439

Escarré J, Lefebvre C, Gruber W, Leblanc M, Lepart J, Riviere Y, Delay B (2000) Zinc and cadmium hyperaccumulation by *Thlaspi caerulescens* from metalliferous and nonmetalliferous sites in the Mediterranean area: implications for phytoremediation. New Phytologist 145, 429–437

Filatov V, Dowdle J, Smirnoff N, Ford-Lloyd B, Newbury HJ, Macnair MR (2006) Comparison of gene expression in segregating families identifies genes and genomic regions involved in a novel adaptation, zinc hyperaccumulation. Molecular Ecology 15, 3045-3059

Filatov V, Dowdle J, Smirnoff N, Ford-Lloyd B, Newbury HJ, Macnair MR (2007) A quantitative trait loci analysis of Zinc hyperaccumulation in *Arabidopsis halleri*. New Phytologist 174, 580-590

Frérot H, Petit C, Lefèbvre C, Gruber W, Collin C, Escarré J (2003) Zinc and cadmium accumulation in controlled crosses between metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* (Brassicaceae). New Phytologist 157, 643–648

Frérot H, Lefèbvre C, Petit C, Collin C, Dos Santos A, Escarré J (2005) Zinc tolerance and hyperaccumulation in F_1 and F_2 offspring from intra and inter-ecotype crosses of *Thlaspi caerulescens*. New Phytologist 165, 111–119

Grotz N, Fox TC, Connolly E, Park W, Guerinot ML, Eide D (1998) Identification of a family of zinc transporters genes from *Arabidopsis* that respond to zinc deficiency. Proceedings of the National Academy of Sciences USA 95, 7220-7224

Hanikenne M, Talke IN, Haydon MJ, Lanz C, Nolte A, Motte P, Kroymann J, Weigel D, Krämer U (2008) Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. Nature 453, 391-395

Johnstone JS, Pepper AE, Hall AE, Chen ZJ, Hodnett G, Drabek J, Lopez R, Price HJ (2005) Evolution of genome size in Brassicaceae. Annals of Botany 95, 229-235

Kearsey MJ and Farquhar AG (1998) QTL analysis in plants; where are we now? Heredity 80, 137-142

Koch MA and Kiefer M (2005) Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of the three diploid species – *Capsella rubella*, *Arabidopsis lyrata* subsp *Petrea*, and *A. thaliana*. Am. J. Bot. 92, 761-767

Kosambi D (1943) The estimation of map distances from the recombination values. Annals of Eugenics 12, 172-175

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M, Mitchell-Olds T, Langley C, Savolainen O (2004) Comparing the

linkage maps of the close relatives *Arabidopsis lyrata* and *Arabidopsis thaliana*. Genetics 168, 1575– 1584

Lombi E, Zhao FJ, Dunham SJ, McGrath SP (2000) Cadmium accumulation in populations of *Thlaspi caerulescens* and *Thlaspi goesingense*. New Phytologist 145, 11–20

Lysak M, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I (2006) Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related species. Proceedings of the National Academy of Sciences USA 103, 5224-5229

Meerts P, Van Isacker N (1997) Heavy metal tolerance and accumulation in metallicolous and non-metallicolous populations of *Thlaspi caerulescens* from continental Europe. Plant Ecology 133, 221–231

Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends in Plant Science 12, 106-117

Peer WA, Mahmoudian M, Freeman JL, Lahner B, Richards EL, Reeves RD, Murphy AS, Salt DE (2006) Assessment of plants from the Brassicaceae family as genetic models for the study of nickel and zinc hyperaccumulation. New Phytologist 172, 248-260

Pence NS, Larsen PB, Ebbs SD, Letham DLD, Lasat MM, Garvin DF, Eide D, Kochian LV (2000) The molecular physiology of heavy metal transport in de Zn/Cd hyperaccumulator *Thlaspi caerulescens*. Proceedings of the National Academy of Sciences USA 97, 4956-4960

Petit JM, van Wuytswinkel O, Briat JF, Lobreaux S (2001) Characterization of an iron-dependent regulatory sequence involved in the transcriptional control of *AtFer1* and *ZmFer1* plant ferritin genes by iron. J Biol Chem. 276, 5584-5590

Qi X, Lindhout P (1997) Development of AFLP markers in barley. Mol Genet Genomics 254, 330-336

Reeves RD, Brooks RR (1983) European species of Thlaspi L. (Crucifereae) as indicators of nickel and zinc. Journal of Geochemical Exploration 18, 275–283

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43, 223-2225

Richau KH, Schat H (2009) Intraspecific variation of nickel and zinc accumulation and tolerance in the hyperaccumulator *Thlaspi caerulescens*. Plant and Soil 314, 253-262

Rigola, D, Fiers M, Vurro E, Aarts MGM (2006) The heavy metal hyperaccumulator *Thlaspi caerulescens* expresses many species-specific genes as identified by comparative EST analysis. New Phytologist 170, 753-766

Roosens N, Verbruggen N, Meerts P, Ximenez-Embun P, Smith JAC (2003) Natural variation in cadmium tolerance and its relationship to metal hyperaccumulation for seven populations of Thlaspi *caerulescens* from western Europe. Plant, Cell and Environment 26, 1657–1672

Schaaf G, Honsbein A, Meda AR, Kirchner S, Wipf D, von Wiren N (2006) *AtIREG2* Encodes a Tonoplast Transport Protein Involved in

Iron-dependent Nickel Detoxification in *Arabidopsis thaliana* Roots. J. Biol. Chem 281, 25532-25540

Schranz E, Lysak MA, Mitchell-Olds T (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer. Trends in Plant Science 11, 535-542

Shen ZG, Zhao FJ, McGrath SP (1997) Uptake and transport of zinc in the hyperaccumulator *Thlaspi caerulescens* and the non-hyperaccumulator *Thlaspi ochroleucum*. Plant, Cell and Environment 20, 898–906

Stam P (1993) Construction of integrated linkage maps by means of a new computer package: JOINMAP. Plant Journal 3, 739-744

Stam P, van Ooijen JW (1996) JOINMAP version 3.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen

Thomine S, Wang RC, Ward JM, Crawford NM, Schroeder JI (2003) cadmium and iron transport by members of a plant metal transporter family in Arabidopsis with homology to Nramp genes. Prceedings of the National Academy of Sciences USA 97, 4991-4996

Van Hoof NALM, Koevoets PL, Hakvoort HW, ten Bookum WM, Schat H, Verkleij JAC, Ernst WHO (2001) Enhanced ATP-dependant copper efflux across the root cell plasma membrane in copper-tolerant *Silene vulgaris*. Plant Physiology 113, 225:232

Van de Mortel JE, Almar Villanueva L, Schat H, Kwekkeboom J, Coughlan S, Moerland PD, Ver Loren van Themaat E, Koornneef M, Aarts MGM (2006) Large Expression Differences in Genes for Iron and Zinc Homeostasis, Stress Response, and Lignin

Biosynthesis Distinguish Roots of *Arabidopsis thaliana* and the Related Metal Hyperaccumulator *Thlaspi caerulescens*. Plant Physiology 142, 1127-1147

van Ooijen JW (1999) LOD significant thresholds for QTL analysis in experimental populations of diploid species. Heredity 89, 803-811

van Ooijen JW (2004) MapQTL® 5, software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V., Wageningen, Netherlands

Vert G, Briat JF, Curie C (2001) Arabidopsis IRT2 gene encodes a root-periphery iron transporter. Plant Journal 26, 181-189

Vert G, Grotz N, Dedaldechamp F, Gaymard F, Guerinot ML, Briata JF, Curie C (2002) IRT1, an Arabidopsis transporter essential for iron uptake from the soil and for plant growth. Plant Cell 14, 1223-1233

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M (1995) AFLP®: a new technique for DNA fingerprinting. Nucleic Acids Research 23, 4407-4414

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. Heredity 93, 77-78

Willems G, Dräger DB, Courbot M, Godé C, Verbrugen N, Saumitou-Laprade P (2007) The genetic basis of zinc tolerance in the metallophyte *Arabidopsis halleri* ssp. *halleri* (Brassicaceae): an analysis of quantitative trait loci. Genetics 176, 659 – 674

Xing JP, Jiang RF, Ueno D, Ma JF, Schat H, McGrath SP, Zhao FJ (2008) Variation in root to shoot translocation of cadmium and zinc

among different accessions of the hyperaccumulators *Thlaspi caerulescens* and *Thlaspi praecox*. New Phytologist 178, 315-325

Xu S (2003) Theoretical basis of the Beavis effect. Genetics 165, 2259-2268

Yogeeswaran K, Frary KA, York TL, Amenta A, Lesser AH, Nasrallah JB, Tanksley SD, Nasrallah ME (2005) Comparative genome analyses of Arabidopsis spp.: Inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. Genome research 15, 505-515

Zha HG, Jiang RF, Zhao FJ, Vooijs R, Schat H, Barker JHA, McGrath SP (2004) Co-segregation analysis of cadmium and zinc accumulation in *Thlaspi caerulescens* interecotypic crosses. New Phytologist 163, 299-312

Zhao FJ, Hamon RE, Lombi E, McLaughlin MJ, McGraph SP (2002) Characteristics of cadmium uptake in two contrasting ecotypes of the hyperaccumulator *Thlaspi caerulescens*. Journal of Experimental Botany 53, 535-543

Zhou J and Goldsbrough PB (2005) Structure, organization and expression of the metallothionein gene family in Arabidopsis. Mol. Gen. Gent 248, 318-328

Chapter 4

Genetic variation analysis among 25 European accessions of the heavy metal hyperaccumulator *Thlaspi caerulescens*

Deniau AX, Grillet L, Pirondini A, Hakvoort HVJ, Grispen VMJ,

Schat H and Aarts MGM

Abstract

Identifying the genes involved in heavy metal hyperaccumulation traits and evidencing the footprints of the natural selection acting on them are essential to understand the evolution of the trait. We observed a considerable variation in the degree of heavy metal hyperaccumulation between Thlaspi caerulescens accessions when grown under controlled conditions in shoot and in root, respectively, for three metals (Cd, Zn and Ni). Contrasting with the strong differences in quantitative traits (Cd, Zn and Ni hyperaccumulation in shoot and in root, respectively), we barely found differentiation for metal-related and non-metal-related molecular markers. Phylogeny analysis identified two major geographic groups, i.e. the accessions from northwestern Europe which were all strongly related to each other and the accessions from southern France which were less related among each other. Two isolated serpentine accessions from northern Italy and northern Spain were more distant from these groups as well as from each other. Genetic distances between accessions varied between markers but were overall too low to infer any conclusions on selection. Non-synonymous and synonymous distances to Arabidopsis varied between markers but were not significantly different from each other, except for ZNT1, MRP4, NRAMP3 and NAS4, which showed an excess of synonymous substitutions, suggesting purifying selection. The patterns of genetic variation among accessions were not correlated with their metal accumulation behavior, nor with the degree and kind of soil metal enrichment at the site of origin. Adaptation to metaliferous soils appears to have occurred repeatedly at local scale.

Introduction

Thlaspi caerulescens is a useful model species to unravel the molecular mechanisms of heavy metal tolerance and hyperaccumulation traits (Assunção *et al.,* 2003a; Deniau *et al*., 2007). It is found on a broad range of soil types including serpentine, calamine and non-metaliferous soils (Escarré *et al*., 2000; Roosens *et al.,* 2003; Basic *et al.,* 2006; Deschamps *et al*., 2007; Jiminez-Ambriz *et al*, 2007; Besnard *et al*., 2009). Controlled experiments have revealed a broad, largely uncorrelated variation of metal-specific tolerance and accumulation characteristics among accessions. The metallicolous accessions usually show enhanced levels of tolerance, compared with non-metallicolous ones, specifically with regard to the metals that are toxically enriched in the soil at the population sites (Meerts & Van Isacker, 1997; Escarré *et al*., 2000; Assunção *et al*., 2003b; Richau & Schat, 2009). When exposed to the same metal concentrations in soil or nutrient solution, there appears to be a broad variation in metal accumulation capacity, which may be negatively (e.g., for Zn), or positively (e.g., for Cd or Ni) correlated with the metal concentrations in the soil at the sites of accession origin (Meerts & Van Isacker, 1997; Escarré *et al*., 2000; Assunção *et al*., 2003b; Richau & Schat, 2009). Metal tolerance and metal accumulation capacity, in so far as they segregate within intra-specific crosses, seem to be largely under independent genetic control, irrespective of the metal in question (Assunção *et al*., 2003c; Zha *et al*., 2004; Richau & Schat, 2009).

Although *T. caerulescens* occurs throughout Europe, its distribution pattern is strongly disjunct, consisting of relatively small and widely interspaced local accessions. Therefore, the gene flow rates are supposed to be limited, which would allow a high degree of genetic differentiation among local accessions, possibly accelerated by a high rate of self-fertilization (Riley, 1956) although, on the other hand, some accessions have been demonstrated to exhibit considerable outbreeding rates (Koch *et al*., 1998; Dubois et al., 2003). Jiminez-Ambriz *et al*. (2007) investigated the patterns of microsatellite diversity among a number of neighboring metallicolous and non-metallicolous accessions in the region around Montpellier (South-France). The results were consistent with a polyphyletic origin of the metallicolous accessions, imposed by divergent selection at a small local scale, with gene flow only weakly opposing local adaptation despite geographical proximity. However, for north-western and central Europe, the possibility of relatively recent long-distance dispersal from common ancestral accessions has been suggested (Koch *et al*., 1998). Based on isozyme variation, these authors suggest that heavy metal tolerance might have originated at least twice within the northern part of Europe, i.e. separately in Great Britain and in continental Europe, as was earlier suggested by Garcia-Gonzales & Clark (1989).

In this study we assessed the pattern of genetic variation among 25 *T. caerulescens* accessions from different parts of Europe, using a set of 15 molecular markers, with the aim to shed more light on the genetic diversity patterns among the north-western European accessions, as compared to those among the southern-European ones. We also

established the phylogenetic relationships of the accessions under study. Another aim was to shed light on the phylogenetic origin of the northern metallicolous accessions. We also addressed the issues of a possible phylogenetic bias of metal accumulation characteristics and the possible role of soil metal compositions as determinants of the variation in these characteristics. Finally, we assessed the non-synonymous and synonymous substitution rates in the coding sequences of metal-related markers between *T. caerulescens* and *Arabidopsis thaliana*, to detect possible effects of divergent directional selection on the structure of the encoded proteins since their last common ancestor.

Material & Methods

Plant materials

Seeds of *T. caerulescens* accessions were collected in the field. Five accessions are originating from United Kingdom, three from Finland, one from Luxembourg, one from Belgium, one from Italy, 13 from France and one from Spain (Table 1). The French accessions were sampled from the Cévennes region at metalliferous (M) and nonmetalliferous (NM) sites, which are only few kilometers apart. M accessions were found on abandoned Zn and lead (Pb) mine wastes (Ganges [GA], which originates in fact from a mine near St-Laurent-le-Minier, but is commonly named Ganges in the literature), Les Malines (LM), Moyen-Age (MA), St Bresson (SB), La Ramponèche (RP), Durfort (DF), St Félix de Pallières (SFP), Le Blémard (LB), Col du Mas de l'Ayre (CMA)) and NM accessions were sampled in the vicinity of the M accessions (Col de la Baraquette (CB), Rochers de la Tude (RT)), or in the nearby region at Causse du Larzac (Montagne de Séranne (MS)). Monte Prinzera (MP), originating from northern Italy and Puente

Location	Country	References	Abbreviations	Soil pH	Zn	Pb	Cd	Ni	Cu
Garigill a*	United Kingdom	Baker *et al.*, 1994	GAR	6,9	808	2212	4	-	-
Coldstones Quarry*	United Kingdom	-	CQ	-	-	-	-	-	-
Aberlyn*	United Kingdom	-	AB	-	-	-	-	-	-
Blackmoor*	United Kingdom	-	BL	-	-	-	-	-	-
Bonsall Moor a*	United Kingdom	Ingrouille & Smirnoff, 1986	BM	7,1	519	610	22	-	-
Black Rock a*	United Kingdom	Baker *et al.*, 1994	BR	7,1	3435	2194	34	-	-
Monte Prinzera	Italy	Assunção *et al.*, 2003c	MP	6,5	104,9	72,5	5	1949,8	35,4
Kuopio - Presidentinkatu	Finland	-	TL	7,1	66,6	24	7,7	48,1	24,7
Kuopio - Tiukanlinna	Finland	-	VN	6,0	77,8	26,3	6,8	37,8	19,8
Kuopio - Vainolanniemi	Finland	-	PK	5,5	58,4	22	8,9	41,7	37,3
La Calamine*	Belgium (Benelux)	Deniau *et al.*, 2006	LC	6,3	61638,9	22806	51,5	198,7	52,3
Lellingen	Luxembourg (Benelux)	Assunção *et al.*, 2003c	LE	5,4	256	173,6	7,4	91,5	33,8
Puente Basadre b	Spain	Lloyd-Thomas, 1995	PB	-	72	35	4	2300	20
Ganges*	South of France	Roosens *et al.*, 2003	GA	-	-	-	-	-	-
Col du Mas de l'Ayre*	South of France	Peer *et al.*, 2003	CMA	6,3	4019,3	951,3	29,9	70,4	34,3
Durfort*	South of France	-	DF	7,0	16764,7	2414,7	46,6	55,6	24,7
Le Bleymard*	South of France	Peer *et al.*, 2003	LB	6,6	28410,1	9626,1	114,2	48,8	50,4
Les Malines*	South of France	Peer *et al.*, 2003	LM	6,4	49647,5	14939,5	426,4	64,8	105,6
La Ramponeche*	South of France	-	RP	6,5	4171,1	4967,9	18,4	26,1	73,6
St Felix de Pallieres*	South of France	Peer *et al.*, 2003	SFP	7,0	2294,2	2559,8	16,8	59,2	49,9
St Bresson* c	South of France	Dubois *et al.*, 2005	SB	7,1	2650	3800	23	-	-
Moyen-Age* c	South of France	Dubois *et al.*, 2005	MA	5,9	1250	1660	11,3	-	-
Rochers de la Tude c	South of France	Dubois *et al.*, 2005	RT	-	46,9	67,1	1,4	-	-
Col de la Baraquette c	South of France	Dubois *et al.*, 2005	CB	-	40	52	1,3	-	-
Montagne de Seranne c	South of France	Dubois *et al.*, 2005	MS	7,1	7	5	1,9	-	-

Table 1: Total metal concentrations (mg kg^{-1} dry mass) and pH of soil samples collected of the 25 *Thlaspi caerulescens* accessions; *: zinc/lead mines; -: not determined; a: by Baker *et al.*, 1994; b: by Lloyd-Thomas, 1995 (soil pH was determined in 1:1 (w-v, H2O) samples. Metals were analyzed after digestion of samples in a 1:20 (w-v) mixture with 4:1 (v-v) HCl:HNO3); c: by Dubois *et al.*, 2005 (concentrations are expressed as mg kg^{-1} ammonium acetate-ethylenediaminetetraacetic acid (EDTA) extractable element)

Basadre (PB), originating from Galicia, Spain, are two accessions growing on serpentine soil. The accessions La Calamine (LC, Belgium) and Lellingen (LE, Luxembourg), from calamine and non-metalliferous soil respectively, have been described in Assunção *et al.* 2003b).

Culture conditions

Plants were grown from seeds sown on moist peat. Three-week-old seedlings were transferred to 1-L polyethylene pots (two seedlings per pot) filled with modified half-strength Hoagland's nutrient solution: 3 mM KNO_3, 2 mM $Ca(NO_3)_2$, 1 mM $NH_4H_2PO_4$, 0.5 mM $MgSO_4$, 1 µM KCl, 25 µmol H_3BO_3, 2 µM $ZnSO_4$, 2 µM $MnSO_4$, 0.1 µM $CuSO_4$, 0.1 µM $(NH_4)6Mo_7O_{24}$, 20 µM Fe(Na)ethylene-diamine-tetra-acetic acid (EDTA). The pH buffer 2-(N-morpholino)ethane-sulfonic acid (MES) was added to a final concentration of 2 mM, and the pH was set at 5.5 using KOH. The solutions were replaced twice a week. All the experiments were performed in a climate chamber (20/15°C day/night; photon flux density 250 $\mu mol.m^{-2}.s^{-1}$ at plant level, 14 $h.d^{-1}$; relative humidity 75% RH). To induce flowering, pots with five-week-old plants were transferred to a cold growth cabinet (4°C day and night; 200 $\mu mol.m^{-2}.s^{-1}$ at plant level; 12 $h.d^{-1}$; ±60% RH) where they remained for 5-6 weeks, after which they returned to the climate chamber.

Cd, Zn and Ni accumulation in the different accessions

Three seedlings per accession were grown in non-metal-amended nutrient solution. After one week 5 μM $CdSO_4$ (Cd treatment), or 10 μM $ZnSO_4$ (Zn treatment), or 10 μM $NiSO_4$ (Ni treatment) were added. These concentration levels were found to yield the highest relative differences in Cd, Zn and Ni accumulation between accessions (Assunção *et al.,* 2003b), respectively. The nutrient solutions were replaced by fresh ones twice a week, and after two weeks of exposure shoots and roots were harvested. The roots were desorbed for 30 minutes in 5 mM ice-cold $Pb(NO_3)_2$. Roots and shoots were dried overnight in a stove at 70°C, wet-ashed in a 4:1 mixture of HNO_3 (65%) and HCl (37%) in Teflon bombs at 140°C for 7 h and analyzed for Cd or Zn or Ni using flame atomic absorption spectrometry (Perkin Elmer 1100B, Perkin Elmer Nederland, Nieuwerkerk a/d IJssel, The Netherlands).

Promoter primers design

Dobeš *et al*. (2004) provided the primers set to amplify TrnL/F Intergenic Spacer region and TrnL intron which are commonly used in phylogenic studies (Table 2). The sequences of the promoters chosen to be used as markers were not available, except for the ZNT1 promoter (Talukdar *et al*., unpublished). A procedure based on sequence homology and micro-colinearity between *Arabidopsis thaliana* and *Thlaspi caerulescens* genes was adopted according to Talukdar *et al.* (unpublished).

Overall	Sequences			Ortholog in *A. thaliana*	Primers
	size (in bp)	Coding regions	Non-coding regions		
All data	6715				
metal responsive genes	4276				
non-metal responsive genes	2439				
IRT1	250	1-250		At4g19680	F- CTGCATCCTCCAGGCAGATTTTACAAAT
					R- TGAGTACGAGGAAGATTGTTTTCATGAG
MRP4	251	1-72, 148-245	73-147, 246-251	At3g62700	F- GTTTGGAGATAAGACAGAGATCGGTGG
					R- GCACAATCATCCCATCCCGC
MT2a	787	157-342	1-156, 343-787	At3g09390	F- GGAAACTGTGGTTGTGGATCTGG
					R- ACGGGACCATCAAACTAAGGTCA
MT2b	551	396-543	1-395, 544-551	At5g02380	F- GAGAAAATGTCTTGCTGTGGAGGAAAC
					R- CGCCGGCGAAATCAACCATT
NAS3	257	1-257		At5g04950	F- ATGCTCGTCGAGTCTGCCGTAGGA
					R- TGTCGACATTTTTGGAAGGTTTTAAGC
NAS4	463	1-463		At1g56420	F- GGCTTTGAGCTGATTGATGATCCTTA
					R- CGAGATCTTTTCGTACAGATCGCAGATC
NRAMP3	320	1-118, 226-320	119-225	At2g23150	F- AGAACTACGGAATAAGGAAGCTCGAGG
					R- GAGCTGAGTGAAGAAACACGTTGTGC
NRAMP4	479	1-479		At5g67320	F- AATCGCCGGGTACCGGAAAG
					R- TTACAACTCCAGCCCAGAGAGGAATC
ZNT1	313	1-177, 251-313	178-250	At1g10970	F- CGTGGTTGTGGAAAGAGGGAATG
					R- TGCAATGAGAGGCCTGATCGTG
ZNT2	603	1-259, 517-556	260-516, 557-603	At1g60960	F- CGTAAG ACCCCAATGTTCTTCATCG
					R- AAAGCAGCAACCATGGCGAA
trn L/F intron	395		1-395		F- CGAAATCGGTAGACGCTACG
					R- GGGGATAGAGGGACTTGAAC
trn L/F IGS	324		1-324		F- GGTTCAAGTCCCTCTATCCC
					R- GATTTTCAGTCCTCTGCTCTAC
Prom ZNT1	760		1-760		F- TTTATCATATCAACCGGGTGCTTCTG
					R- TCATGGGAACAAAGAGTGTCTTCTTCTT
Prom NAS3	385		1-385		F- TGTATAATGGTTAAATGACGAATGGTCAA
					R- TCCAACCGTTTATCCCTCCGAA
Prom At5g19440	578		1-578		F- CTCCCCACCAACTTCCATTTACTTAGT
					R- GGACGCTCCTGTTACACACACCACT

Table 2: Description of the metal and non-metal-related

A 5' outward primer (Figure 1A) at the beginning of the known *T. caerulescens* gene sequence was designed. Then, the upstream gene of the orthologous gene was fished out in *A. thaliana* by databank analysis and one primer was designed in the 3' region of this gene (Figure 1B). These primer pairs were used on the *T. caerulescens* genomic DNA to amplify the putative promoter region (Figure 1C).

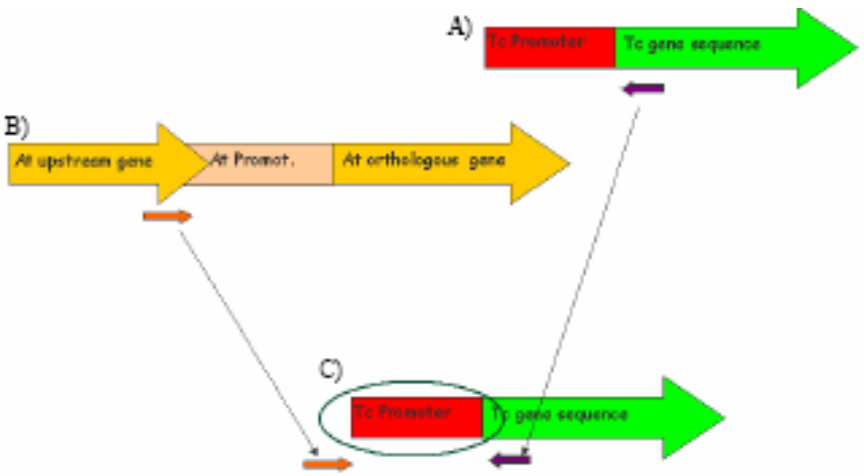


Figure 1: Procedure adopted to obtain a pair of primers for the amplification of the promoter region. Tc = *Thlaspi caerulescens*; At = *Arabidopsis thaliana*

Loci amplification and sequencing

DNA of all accessions was extracted according to Deniau *et al.* (2006). Thlaspi genes or promoters were amplified using the primers described in Table 2, following the PCR program: 94°C for 2 min, followed by 34 cycles (94°C for 30 sec, 58°C for 30sec, 72°C for 1 min) and a final elongation at 72°C for 5 min. Then, 1 µl of PCR products were checked on agarose gel (0.8%). The 19 µl of PCR product remaining was purified

by precipitation. 40 µl H_2O, 5 µl NaAc and 135 µl of 96% EtOH were added to each PCR reaction and were centrifuged for 45 min. at 13.000 rpm. PCR products were cleaned with 70% EtOH and eluted in 20 µl H_2O. A small amount of the eluted solution was checked on agarose gel to quantify DNA. Approximately 25 ng of DNA were used for the ligation reaction following the protocol for ligation "Using the pGEM®-T Easy Vectors and the 2X Rapid Ligation Buffer" (Technical Manual, pGEM®-T Easy Vectors, Promega). The transformation was performed using high-efficiency competent cells (>1·108 cfu/µg DNA). Fifty µl of competent cells and 2 µl of each ligation product were mixed and a potential of 2.5 kV was applied to induce the consequent transformation. 200 µl of SOC medium was added to the transformed competent cells (2% Bacto®-tryptone, 0.5% Bacto®-yeast extract, 10 mM NaCl, 2,5 mM KCl, 20 mM Mg2+ stock, 20 mM glucose) and incubated at 37° C for one hour. LB plates (1% Bacto®-tryptone, 0.5% Bacto®-yeast extract, 0.5% NaCl and 1.2% Agar) containing ampicillin (250 $\mu g.ml^{-1}$), IPTG (16 $\mu g.ml^{-1}$) and X-Gal (32 $\mu g.ml^{-1}$) were prepared. Fifty µl of sample was added and spread over the LB plates and incubated overnight at 37°C. Blue colonies (transformed) were confirmed by PCR for the presence of the right PCR product. Positives colonies were therefore inoculated in 10 ml tubes containing 2 ml of LB liquid medium (1% Bacto®-tryptone, 0.5% Bacto®-yeast extract, 0.5% NaCl), ampicillin (250 $\mu g.ml^{-1}$) and ITPG (16 $\mu g.ml^{-1}$). These tubes were incubated while shaking (200 rpm), at 37° C, for 12 hours. The overnight cell culture was centrifuged and cells were resuspended in 200 µl of resuspension buffer (50mM Tris (pH=8), 10 mM EDTA (pH=8)). Two hundred µl of a 0.2 M

NaOH and 1% SDS buffer was added to the samples. After gently mixing by inversion,200 µl of third buffer (60mL 5M KAc; 11.5mL HAc; 28.5mL H_2O) was finally added. Then, samples were centrifuged at 13000 rpm and supernatant was transferred to a new eppendorf tubes. A phenol/chlorophorm/isoamynoalcohol (400 µl) extraction was performed. The plasmid pellet was precipitated with 96% EtOH, cleaned with 70% EtOH and finally resuspended in 100 µl H_2O containing 100ng/ml RNAse A. A small amount of plasmid was run on agarose gel to estimate sample concentration.

DNA sequences were obtained on both strands using a "Big Dye terminator protocol"; this procedure reduces the possibility of a sequencing mistake. Each reaction mix was prepared adding 10ng of plasmid, 2 µl of Ready reaction premix, 1 µl of Big dye Sequencing buffer (2.5x), 1 µl of 5 $pmol.\mu l^{-1}$ of forward or reverse primer to a final volume of 10 µl of MilliQ water. The amplification of the target regions was performed following the PCR program: 94°C for 2 min, followed by 25 cycles (94°C for 30 sec, 58°C for 30sec, 72°C for 1 min) and a final elongation at 72°C for 5 min. The target loci were sequenced on an ABI3100 sequencing machine (Greenomics, Plant Research International, Wageningen University and Research Centre, The Netherlands) and were analyzed using Vector NTI software package. Two internal primers were designed on the putative promoter sequence available to target the Thlaspi-specific promoter region, preventing the possibility of non-specific amplification due to the use of *A. thaliana* primers. Each PCR fragment was sequenced in both ways (forward and reverse) to reduce sequencing mistakes.

Data analysis

For each fragment the forward and reverse sequences were assembled using ContigExpress Software (component of Vector NTI 10 advanced software package). Both chromatograms were compared and ambiguities were analyzed and corrected. The different consensus sequences were aligned using "AlignX" (component of Vector NTI 10 advanced software package) with additional adjustments by hand in case of ambiguous alignment. The final alignment was exported into a new data file (.emf) and converted by Treecon software (Van de Peer & Wachter, 1994) in order to be analyzed by the phylogenic software MEGA 3.1 (Kumar *et al.*, 2004).

MEGA software permits the visualization of the final alignment as a matrix where homologous sites are shown by dots and only the polymorphic sites are explicitly displayed. Coding regions, introns and other non-coding regions were defined in order to give different weights in the analysis. Evolutionary distances were calculated by MEGA using two different approaches such as nucleotide distances and synonymous and non-synonymous distances. Nucleotide distances estimate the number of nucleotide substitutions per site among DNA sequences, it was calculated according to the Kimura's two parameter model (K2P) (Kimura, 1980).

Results

Cd, Zn and Ni accumulation in different accessions

The different Thlaspi accessions showed a large variation of Cd accumulation in shoot (46,3-fold difference) and in root (22,9-fold difference) respectively, with no obvious correlation between root and shoot accumulation (Figure 2).

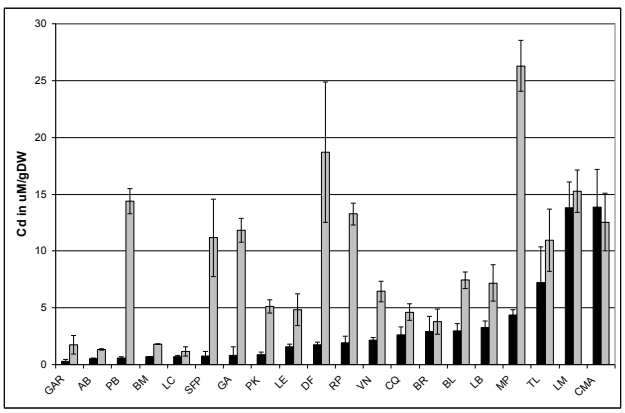


Figure 2: Cd accumulation in 20 different *Thlaspi caerulescens* accessions in shoot (black bars) and in root (grey bars) and their standard error, respectively. Accessions are ranked by their increasing Cd accumulation in shoot.

The accessions LM and CMA, both from southern France, accumulated Cd to much higher concentrations in their shoots than any of the others. These accessions also showed the high shoot to root Cd concentration ratios (≥ 1) considered to be typical of hyperaccumulation behaviour, whereas the other accessions usually exhibited much higher concentrations in roots than in shoots, One of the Finnish accessions, TL (non-metallicolous), as well as the Italian serpentine accession, MP,

showed relatively high shoot Cd concentrations too. Cd accumulation in the root was particularly high in all of the southern-French accessions tested, but also in one of the Finnish ones (TL), and in the two serpentine accessions, from Italy and Spain, respectively (MP and PB).

As for Cd, the accessions showed a broad variation in Zn accumulation in shoot and in root, again without correlation between root and shoot accumulation (Figure 3).

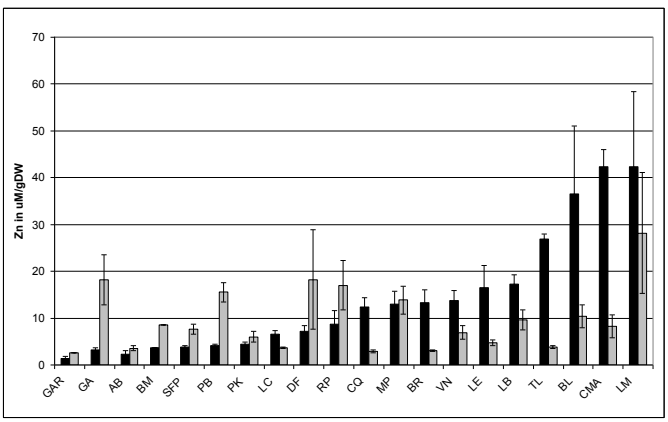


Figure 3: Zn accumulation in 20 different *Thlaspi caerulescens* accessions in shoot (black bars) and in root (grey bars) and their standard error, respectively. Accessions are ranked by their increasing Zn accumulation in shoot.

As for Cd, LM and CMA showed the highest Zn accumulation in shoot, followed by BL (England) and TL (Finland), which were also good shoot Cd accumulators. As for Cd, the Zn concentrations in roots were high in LM, RP, DF, GA (southern France), and in the serpentine accessions MP and PB. Among the ten accessions with the highest shoot Zn accumulation, nine had shoot to root Zn concentration ratios above unity. On the other hand, nine out of ten of the accessions with the

lowest shoot Zn accumulation had shoot to root concentration ratios lower than one. Big differences were also found for Ni accumulation in shoot and root (Figure 4).

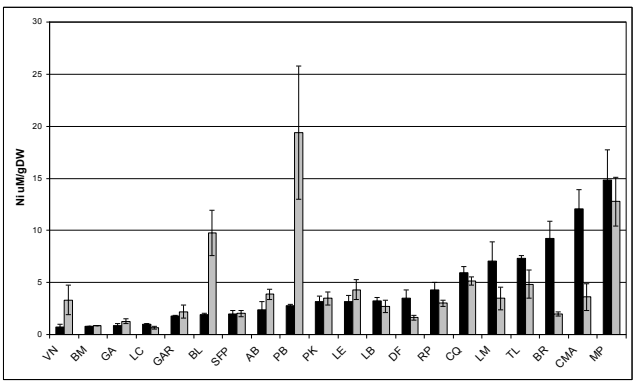


Figure 4: Ni accumulation in 20 different *Thlaspi caerulescens* accessions in shoot (black bars) and in root (grey bars) and their standard error, respectively. Accessions are ranked by their increasing Ni accumulation in shoot.

MP is the highest shoot Ni accumulator, followed by CMA, BR, TL and LM, of which all but BR (England) were also among the best Cd and Zn shoot accumulators. High root Ni accumulation was apparent in PB and BL, which were also among the higher root Cd and Zn accumulators. Like in case of Zn, out of the ten accessions with the highest shoot Ni accumulation, nine had a shoot to root Ni concentration ratio above one, whereas out of the ten accessions with the poorest shoot Ni accumulation, nine had a shoot to root Ni concentration ratio below unity.

Phylogeny analysis

A 50%-majority-rule consensus tree was generated using the maximum parsimony method by the combination of all the data obtained after alignment (Figure 5).

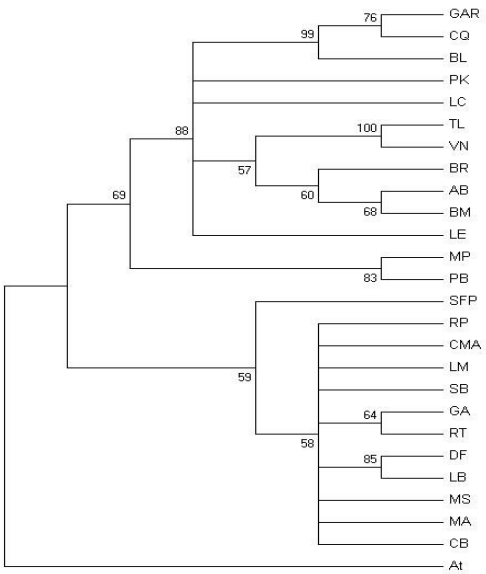


Figure 5: 50%-majority-rule consensus tree for *Thlaspi caerulescens* from different sites in Europe. The consensus tree was generated by the combination of all the data achieved using the Maximum Parsimony method (1000 bootstrap replicates)

The different accessions are grouped in separate clusters. In particular, there are two main clusters: one grouping all the accessions from southern France, and the other one grouping the remaining accessions. Among the southern France accessions SFP seemed to be slightly different from the rest of the accessions originating from this region. In the other cluster, it is clear that the accessions from United Kingdom, Finland and Benelux are clustering together. Also, MP and PB, both originating from serpentine soils are grouped together.

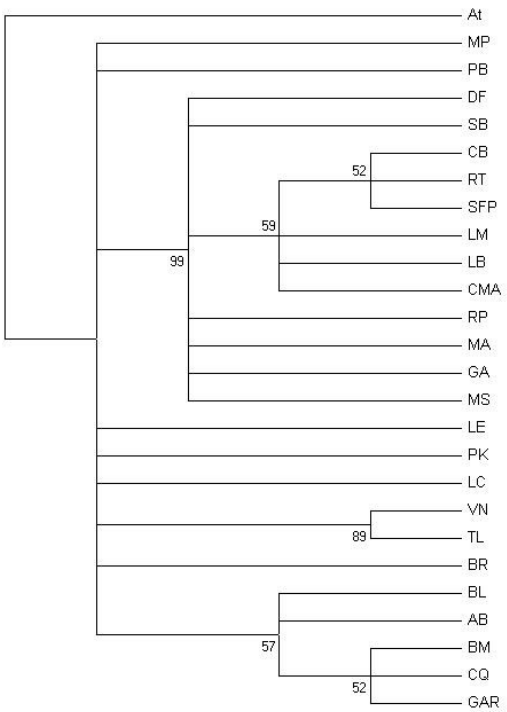


Figure 6: 50%-majority-rule consensus tree for *Thlaspi caerulescens* from different sites in Europe. The consensus tree was generated by the combination of all coding data achieved using the Maximum Parsimony method (1000 bootstrap replicates)

In Figure 6, only the coding regions were used. Again all the accessions from southern France stood apart from the rest of the accessions. This latter clade appeared in 99% of the trees calculated. In this tree, the accessions from United Kingdom grouped together, except for BR. Another 50%-majority-rule consensus tree has been calculated using the non-coding sequences available (Figure 7).

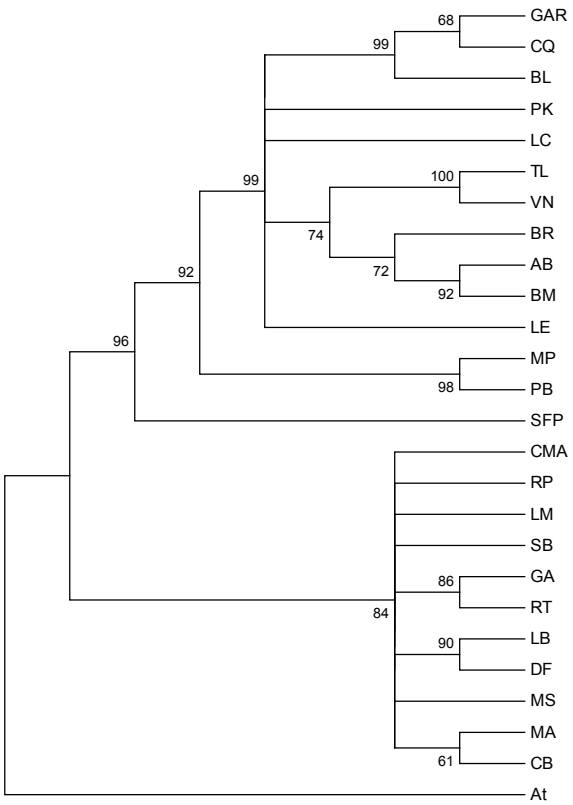


Figure 7: 50%-majority-rule consensus tree for *Thlaspi caerulescens* from different sites in Europe. The consensus tree was generated by the combination of all non-coding data achieved using the Maximum Parsimony method (1000 bootstrap replicates)

As for the two other trees calculated, the accessions originating from southern France are clearly in a well-supported clade (bootstrap value of 85). Remarkably, SFP is not included in this clade but it is genetically closer to the accessions form regions other rather than southern France. This tree shows again a large similarity between the accessions from United Kingdom, Finland and Benelux. Within the southern-French clade, the metallicolous accessions do not cluster. Within the north-western European group, the British accessions, which are all

metallicolous are grouped into two clusters which are no more related to the continental metallicolous population, LC, then to the continental non-metallicolous accessions.

Within the non-metal related markers trnL/F intron and trnL/F IGS, we surprisingly found only one parsimony-informative site out of a total of 719. The genetic distances between the *Thlaspi caerulescens* accessions were overall small (Table 3), especially for the non-metal related markers. The largest distances were found for MT2b (0,0560) and the NAS3 promoter (0,0513). The distances between the non-coding regions were many-fold higher than those between the coding regions, in particular for NRAMP3 which did not show any substitutions in the coding region. For all the coding sequences, the non-synonymous distances were not significantly different from the synonymous ones. We also calculated the mean genetic distances within and among the different geographical groups (Table 4). Based on these data, a clear pattern emerges with 4 major groups, i.e. the accessions from southern France, those from northwestern Europe (United-Kingdom, Benelux and Finland) and the 2 isolated accessions from Spain and Italy. The group from southern France is the most distant from all the others. Moreover, this group shows the largest mutual distances between the members. The Thlaspi accessions from United Kingdom, Finland and Benelux seemed to be strongly related (distance ≤ 0.0071).

	Genetic distance among Tc accessions			Genetic distance between Tc accessions and At		
	All sequences	coding	non-coding	All sequences	coding	non-coding
All data	0,0176 (0,0010)	0,0069 (0,0010)	0,0255 (0,0015)	0,1996 (0,0071)	0,1140 (0,0064)	0,4260 (0,0211)
All metal-responsive genes	0,0182 (0,0013)	0,0069 (0,0010)	0,0405 (0,0032)	0,1996 (0,0077)	0,1140 (0,0065)	0,4260 (0,0224)
All non-metal-responsive genes	0,0165 (0,0017)	NA	0,0165 (0,0017)	NA	NA	NA
IRT1	0,0071 (0,0025)	0,0071 (0,0025)	NA	0,1157 (0,0020)	0,1152	NA
MRP4	0,0073 (0,0040)	0,0041 (0,0031)	0,0146 (0,0074)	0,1229 (0,0237)	0,0724 (0,0210)	0,2584 (0,0668)
MT2a	0,0192 (0,0030)	0,0066 (0,0033)	0,0238 (0,0039)	0,3436 (0,0252)	0,1267 (0,0278)	0,4559 (0,0406)
MT2b	0,0560 (0,0070)	0,0074 (0,0038)	0,0776 (0,0102)	0,3574 (0,0313)	0,0554 (0,0189)	0,5442 (0,0526)
NAS3	0,0041 (0,0023)	0,0041 (0,0023)	NA	0,0646 (0,0156)	0,0646 (0,0156)	NA
NAS4	0,0068 (0,0021)	0,0068 (0,0021)	NA	0,0975 (0,0149)	0,0975 (0,0149)	NA
NRAMP3	0,0108 (0,0030)	0,0008 (0,0005)	0,0352 (0,0106)	0,1377 (0,0211)	0,0645 (0,0174)	0,3528 (0,0750)
NRAMP4	0,0119 (0,0032)	0,0119 (0,0032)	NA	0,1337 (0,0167)	0,1337 (0,0167)	NA
ZNT1	0,0143 (0,0039)	0,0023 (0,0013)	0,0548 (0,0177)	0,3075 (0,0367)	0,2403 (0,0393)	0,5639 (0,1268)
ZNT2	0,0218 (0,0035)	0,0108 (0,0034)	0,0333 (0,0066)	0,2037 (0,0209)	0,1503 (0,0243)	0,2795 (0,0386)
tm_L/F_intron	0,0028 (0,0009)	NA	0,0028 (0,0009)	NA	NA	NA
tm_L/F_IGS	0,0027 (0,0008)	NA	0,0027 (0,0008)	NA	NA	NA
Promoter_ZNT1	0,0196 (0,0033)	NA	0,0196 (0,0033)	NA	NA	NA
Promoter_NAS3	0,0513 (0,0073)	NA	0,0513 (0,0073)	NA	NA	NA
Promoter_At5g1944	0,0091 (0,0022)	NA	0,0091 (0,0022)	NA	NA	NA

Table 3: Genetic distances among *Thlaspi caerulescens* accessions and between *Thlaspi caerulescens* and *Arabidopsis thaliana*. Standard errors are indicated in brackets. NA = Not Applicable.

	United Kingdom	Finland	Benelux	Spain	Italy	South of France
South of France	0,0211 (0,0016)	0,0214 (0,0015)	0,0217 (0,0016)	0,0217 (0,0016)	0,0228 (0,0016)	0,0107 (0,0007)
Italy	0,0165 (0,0016)	0,0172 (0,0016)	0,0165 (0,0014)	0,0138 (0,0015)	NA	
Spain	0,0158 (0,0017)	0,0145 (0,0014)	0,0143 (0,0014)	NA		
Benelux	0,0079 (0,0001)	0,0077 (0,0009)	0,0071 (0,0010)			
Finland	0,0077 (0,0009)	0,0047 (0,0007)				
United Kingdom	0,0048 (0,0006)					

Table 4: Genetic distances within and between different geographical groups of *Thlaspi caerulescens* accessions. Standard errors are indicated in brackets. NA = Not Applicable.

This whole group is rather homogeneous, with distances within and among subgroups consistently less than 0.01. The mean genetic distances within the subgroups are only slightly and insignificantly lower than the average distances between subgroups. The group from southern France is approximately equally distant from all the other groups and the distances within this group are about half of their distances to the other groups. Although PB and MP group together in Figure 5, they are almost equally distant from each other as from the others groups, except for the group from southern France, which is slightly more distant. This pattern was found for the coding regions and the non-coding regions, although the average distances in coding regions and the non-coding regions are different. Within the geographic groups, the genetic distances between the non-metaliferous accessions are not significantly different from their average distances to the metallicolous accessions (data not shown).

Discussion

Cd, Zn and Ni accumulation in different accessions

We observed a considerable variation in the degree of heavy metal hyperaccumulation of the three metals in shoot and in root between accessions when grown under controlled conditions, in line with the previous studies from Assunção *et al.* (2003b) and Peer *et al,* (2003, 2006). However, results will vary in different experimental setups (metal concentration, time of exposure, single or multiple metal exposure, etc.) as shown by Assunção *et al.* (2001, 2003b, 2009). Except for CMA and LM, all the accessions accumulate Cd to significantly higher concentrations in their root than in their shoot under our experimental conditions, confirming that Cd hyperaccumulation may not be a species-wide phenomenon in *Thlaspi caerulescens* (Assunção *et al*., 2003b). On average, the southern-French accessions exhibit higher Cd concentrations, either in root or in shoot, then the north-western European ones. However, some of the latter accessions (TL, BL, BR) show in fact higher Cd concentrations, either in shoot, or both in root and shoot, then the majority of southern-French ones, indicating that Cd hyperaccumulation capacity may not be a strictly southern-French phenomenon. GA is the best studied Cd hyperaccumulator (Lombi et al., 2001; Zhao *et al.*, 2002; Assunção *et al*., 2003a, 2003b; Zha *et al.*, 2004; Deniau *et al.*, 2006). In our experiments, GA is a high Cd root hyperaccumulator but it has a rather low translocation to the shoot, in agreement with the results obtain by Xing *et al.* (2008). In this study, CMA and LM, both from South France, had the highest translocation

rates for Cd, but were also among the better Zn and Ni translocators. Our results clearly suggest a degree of correlation between Cd and Zn translocation: the 10 best Zn translocating accessions were also the 10 best Cd hyperaccumulating accessions. This is in agreement with Xing *et al*. (2008), who found that Cd and Zn translocation capacities were significantly positively correlated among a large number of accessions of *T. caerulescens* and other hyperaccumulating Thlaspi species. The two serpentine accessions, MP and PB, show a remarkably different translocation behavior for Ni. Whereas MP was the best Ni shoot accumulator among the accessions tested, PB was not even among the 10 best ones. On the other hand, PB was the best Ni root accumulator, closely followed by MP. PB was also among the worst translocators of Zn and, particularly, Cd. Overall, our results suggest that the variation in accumulation and translocation behavior among the accessions under study is rather poorly correlated with geographical origin or soil metal composition.

Phylogeny analysis

Contrasting with the strong differences in quantitative traits (Cd, Zn and Ni hyperaccumulation in shoot and in root, respectively), we found only little molecular differentiation between accessions, both for non-metal- and metal-related markers. Chloroplast DNA (cpDNA) markers, which are considered to provide a more conservative reflection of history than nuclear markers, as a result of maternal inheritance and a lack of recombination (Birky *et al.*, 1989; Ennos, 1994; Martinez *et al*., 1997), yielded ony one informative site, out of a total of more than 700. This

result indicates a relatively recent divergence of the *Thlaspi caerulescens* accessions under investigation. Also the nuclear markers yielded fairly little informative variation on average, which was nevertheless sufficient to generate a clear pattern. The north-western European accessions are very closely related. Assuming that the gene flow rates between these accessions are generally low, this would mean that they all must have recently descended from a common ancestral population that colonized the region after the last glacial. The southern European accessions are significantly more distant among each other, suggesting that they diverged much earlier than the northern-European ones. This is already evident within the relatively small geographical area of the southern-French accessions, where the average genetic distance among accessions is twice as high as that within the north-western European subgroups, and even significantly higher than the distances between these subgroups, on average.

There seems to be no obvious phylogenetic bias of the metal-specific accumulation capacities, possibly except for the relatively high Cd accumulation capacity of all the southern-French accessions. It should be noted however, that some accessions from other regions have Cd accumulation capacities within the range of the southern-French ones. In addition, our results are in line with the hypothesis of a multiple origin of the metallicolous accessions. Within the southern-French region there is clearly no grouping according to soil type and the genetic distances within the metallicolous group of accessions are neither significantly different from their average distances to the non-metallicolous accessions, nor from the average distances among the non-metallicolous

accessions. In the north-western European group of accessions there seems to some degree of clustering according to soil type. However, this is only because all the sampled accessions from UK are metallicolous and all the accessions from Finland are non-metallicolous. However, in terms of genetic distance, the continental metallicolous accession LC is on average at least as distant from the metallicolous British accessions (0.0076), as it is from the Finnish non-metallicolous accessions (0.0071) and the neighboring non-metallicolous accession LE (0.0071). In addition, LE is as distant from the non-metallicolous Finnish accessions (0.0085), as it is from the metallicolous British ones (0.0087). The clustering of the two serpentine accessions can neither be taken as evidence of common descent from a serpentine ancestor, since we were not able to sample nearby non-metallicolous accessions. Moreover, MP and PB are in fact rather distantly related. Overall, our results are well compatible with repeated adaptation to metalliferous soil at a local scale, which is also plausible for *A. halleri* too (Pauwels et al, 2005; Meyer *et al*., 2010). Analyses of molecular variation in the facultative metallophytes *Armeria maritima* (Vekemans & Lefèbvre, 1997) and *Silene paradoxa* (Mengoni *et al*., 2001) also suggested recent divergence and multiple evolution of metallicolous accessions at a local scale.

The nuclear markers which have been supposed to be involved in heavy metal hyperaccumulation show a low level of differentiation, except for MT2b and the NAS3 promoter, which indicates that divergent selection might have acted upon these sequences. However, the phylogenetic trees made for these sequences alone were very much like those made for the total of markers, and also in these trees there was no grouping according

to metal accumulation behavior, nor to soil type, indicating that selective pressure by metals, if any, apparently did not affect the tree structure. In order to get any other evidence of selection, we also calculated the synonymous and non-synonymous distances for the coding sequences of the metal-related markers together, as well as for the individual markers. On average, the synonymous distances were slightly but insignificantly longer than the non-synonymous ones, which suggest that at least most of the variation in these markers is selectively neutral. For some individual markers the opposite pattern was found, but statistically significant differences were never found, which is not surprising in view of the low number of substitutions per marker. In conclusion, we did not find any evidence of directional selection on either of these markers. In contrast, one marker, NRAMP3, was without any substitutions, which argues in favor of purifying selection.

Finally we also compared all the *T. caerulescens* markers with the corresponding sequences in *Arabidopsis thaliana*. The overall genetic distance from Arabidopsis was 0.20 (0.11 and 0.43 for coding and non-coding sequences, respectively), with large differences between markers. Of the coding sequences of the metal-related markers, MRP4, MT2b, NAS3 and NRAMP3 were relatively conserved (0.05 – 0.07), whereas, in particular, ZNT1 was strongly diverged (0.24). However, although ZNT1 might be a crucial gene in hyperaccumulation indeed (Pence et al., 2000), the high degree of divergence of this gene is largely due to synonymous substitution. In fact, for this gene the synonymous distance greatly exceeds the non-synonymous one, which is, as such, an indication of predominantly purifying selection, Also for MRP4, NAS4

and NRAMP4, the synonymous distances significantly exceeded the non-synonymous ones, whereas only for one marker, MT2b, the opposite was found. However, this was solely due to a strongly decreased rate of synonymous substitutions, whereas the non-synonymous ones were in fact lower than average. This could be due to the high abundance of the MT2b protein, which could lead to selection against codon bias. Thus, the selection on the coding sequences of the metal-related markers used in this study seems to have been purfying, rather than divergent. This is in agreement with the hypothesis that hyperaccumulation is generally due to changes in gene expression patterns, rather than changes in protein structure (Lasat et al., 1996; van de Mortel et al., 2006). However, although the metal-related genes used in this study are all much higher expressed in *Thlaspi caerulescens* than in *Arabidopsis thaliana* (van de Mortel *et al.*, 2006), it is uncertain whether the phenotypic intraspecific variation among *T. caerulescens* accessions is related to variation in the expression of these genes. Some studies suggest that this may not be the case in general (Assunção *et al*., 2001; Xing *et al.*, 2008).

Acknowledgements

We thank V. Hassinen for collecting the three accessions from Finland on your request. We are grateful to S. Dubois for sharing DNA from the accessions Moyen-Age, St Bresson, Col de la Baraquette, Rochers de la Tude and Montagne de Séranne. Also, the authors thank A. Smith for providing us with seeds from the accessions originating from United Kingdom, as well as N. Verbruggen for the Puente Basadre seeds. This research was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek-Genomics (NWO) Grant (050-10-166).

References

Assunção AGL, Da Costa Marins P, De Folter S, Vooijs R, Schat H, Aarts MGM (2001) Elevated expression of metal transporter genes in three accessions of metal hyperaccumulator *Thlaspi caerulescens*. Plant, Cell and Environment 24: 217 - 226

Assunção AGL, Schat H, Aarts MGM (2003a) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. New phytologist 159: 411 – 419

Assunção AGL, Ten Bookum WM, Nelissen HJM, Vooijs R, Schat H, Ernst WHO (2003b) Differential metal-specific tolerance and accumulation patterns among *Thlaspi caerulescens* populations originating from different soil types. New Phytologist 159, 411-419

Assunção AGL, Ten Bookum WM, Nelissen HJM, Vooijs R, Schat H, Ernst WHO (2003c) A cosegregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator *Thlaspi caerulescens*. New Phytologist 159, 383-390

Assunção AGL, Pieper B, Vromans J, Lindhout P, Aarts MGM, Schat H (2006) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. New Phytologist 170, 21-32

Baker AJM, Reeves RD, Hajar ASM (1994) Heavy metal accumulation and tolerance in British populations of the metallophyte *Thlaspi caerulescens* J. & C. Presl (Brassicaceae). New Phytologist: 127:61-68

Basic N, Salamin N, Keller C, Galland N, Besnard G (2006) Cadmium hyperaccumulation and genetic differentiation of *Thlaspi caerulescens* populations. Biochemistry Systematics and Ecology 9, 667-677

Besnard G, Basic N, Christin PA, Savova-Bianchi D, Galland N (2009) *Thlaspi caerulescens* (Brassicaceae) population genetics in western Switzerland: is the genetic structure affected by natural variation of soil heavy metal concentrations? New Phytologist 181, 974-984

Birky CM, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics, 121:613-627

Dechamps C, Lefèbvre C, Noret N, Meerts P (2007) Reaction norms of life history traits in response to zinc in Thlaspi caerulescens from metalliferous and nonmetalliferous sites. New Phytologist 173, 191-198

Deniau AX, Pieper B, Ten Bookum WM, Lindhout P, Aarts MGM, Schat H (2006) QTL analysis of cadmium and zinc accumulation in the heavy metal hyperaccumulator *Thlaspi caerulescens*. Theoretical and Applied Genetics 113, 907–920

Deniau AX, Schat H, Aarts MGM (2007) Genetics and genomics of the heavy metal hyperaccumulator model species *Thlaspi caerulescens*. Genes, Genomes and Genomics 1, 81-88

Dobeš CH, Mitchell-Olds T, Koch MA (2004) Extensive chloroplast haplotype indicates Pleistocene hybridization and radiation of North

American *Arabis drummondii*, *A. x divaricarpa*, and *A. holboellii* (Brassicaceae). Molecular Ecology 13, 349-370

Dubois S, Cheptou PO, Petit C, Meerts P, Poncelet M, Vekemans X, Lefèbvre C, Escarré J (2003) Genetic structure and mating systems of metallicolous and nonmetallicolous populations of *Thlaspi caerulescens*. New Phytologist 157: 633-641.

Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plant populations. Heredity, 72:250-259.

Escarré J, Lefèbvre C, Gruber W, Leblanc M, Lepart J, Rivière Y, Delay B (2000) Zinc and cadmium hyperaccumulation by *Thlaspi caerulescens* from metalliferous and nonmetalliferous sites in the Mediterranean area: implications for phytoremediation. New Phytologist 145: 429–437

Garzia-Gonzales A, Clark SC (1989) The distribution of *Minuartia verna* and *Thlaspi alpestre* in the British Isles in relation to 13 soil metals. Vegetatio 84, 87-98.

Ingrouille MJ, Smirnoff N (1986) *Thlaspi caerulescens* J & C Presl (*T. alpestre* L.) in Britain. New Phytologist 102, 219-233

Jiménez-Ambriz G, Petit C, Bourrié I, Dubois S, Olivieri I, Ronce O (2007) Life history variation in the heavy metal tolerant plant *Thlaspi caerulescens* growing in a network of contaminanted and noncontaminated sites in southern France: role of gene flow, selection and phenotypic plasticity. New Phytologist 173, 199-215

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16, 111-120

Koch M, Mummenhoff K, Hurka H (1998) Systematics and evolutionary history of heavy metal tolerant *Thlaspi caerulescens* in Western Europe: evidence from genetic studies based on isozyme analysis. Biochemical Systematics and Ecology 26, 823-838

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetic analysis and sequence alignment. Briefings in bioinformatics. 2, 150-163

Lasat MM, Baker AJM, Kochian LV (1998) Physiological characterization of root Zn^{2+} absorption and translocation to shoots in Zn hyperaccumulator and nonaccumulator species of Thlaspi. Plant Physiology 112, 1715-1722

Lloyd-Thomas DH (1995) Heavy Metal Hyperaccumulation by *Thlaspi caerulescens* J. & C. Presl. PhD Thesis. University of Sheffield, Sheffield, UK.

Lombi E, Zhao FJ, McGrath SP, Young SD, Sacchi GA (2001) Physiological evidence for a high-affinity cadmiuym transporter highly expressed in a Thlaspi caerulescens ecotype. New Phytologist 149, 53-60

Martinez P, Lopez C, Roldan M, Sabater B, Martin M (1997) Plastid DNA of five accessions of *Arabidopsis thaliana*: sequence of ndhG gene and maternal inheritance. Plant Science, 123:113-122

Meerts P, Van Isacker N (1997) Heavy metal tolerance and accumulation in metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* from Continental Europe. Plant Ecology 133, 221-231

Mengoni A, Barabesi C, Gonnelli C, Galardi F, Gabbrielli R, Bazzicalupo M (2001) Genetic diversity of heavy metal-tolerant populations in *Silene paradoxa* L. (Caryophyllaceae): a chloroplast microsatellite analysis. Molecular Ecology, 10:1909-1916

Meyer C-L, Kostecka AA, Saumitou-Laprade P, Créach A, Castric V, Pauwels M, Frérot H (2010) Variability of zinc tolerance among and within populations of the pseudometallophyte species *Arabidopsis halleri* and possible role of directional selection. New Phytologist 185: 130-142

Pauwels M, Saumitou-Laprade P, Holl AC, Petit D, Bonnin I (2005) Multiple origin of metallicolous populations of the pseudometallophyte *Arabidopsis halleri* (Brassicaceae) in central Europe: the cpDNA testimony. Molecular Ecology, 14:4403-4414

Peer WA, Mamaoudian M, Lahner B, Reeves RD, Murphy AS, Salt DE (2003) Identifying model metal hyperaccumulating plants: germplasm analysis of 20 Brassicaceae accessions from a wide geographical area. New Phytologist 159, 421-430

Peer WA, Mahmoudian M, Freeman JL, Lahner B, Richards EL, Reeves RD, Murphy AS, Salt DE (2006) Assessment of plants from the Brassicaceae family as genetic models for the study of nickel and zinc hyperaccumulation. New Phytologist 172, 248-260

Pence NS, Larsen PB, Ebbs SD, Letham DLD, Lasat MM, Garvin DF, Eide D, Kochian LV (2000) The molecular physiology of heavy metal transport in de Zn/Cd hyperaccumulator *Thlaspi caerulescens*. Proceedings of the National Academy of Sciences USA 97, 4956-4960

Richau KH, Schat H (2008) Intraspecific variation of nickel and zinc accumulation and tolerance in the hyperaccumulator *Thlaspi caerulescens*. Plant and Soil 314, 253-262

Riley R (1956) The influence of the breeding system on the genecology of *Thlaspi alpestre* L. New Phytol. 55, 319-330

Roosens N, Verbruggen N, Meerts P, Ximinez-Embun P, Smith JAC (2003) Natural variation in cadmium tolerance and its relationship to metal hyperaccumulation for seven populations of *Thlaspi caerulescens* from western Europe. Plant, Cell and Environment 26, 1657-1672

Thomine S, Lelievre F, Debarbieux E, Schroeder JI, Barbier-Brygoo (2003) *AtNRAMP3*, a multispecific vacuolar metal transporter involved in plant responses to iron deficiency. Plant Journal 34, 685-695

Van de Mortel JE, Villanueva LA, Schat H, kwekkeboom J, Coughlan S, Moerland PD, Ver Loren van Themaat E, Koornneef M, Aarts MGM (2006) Large expression differences in genes for iron and zinc homeostasis, stress response, and lignin biosynthesis distinguish roots of *Arabidopsis thaliana* and the related metal hyperaccumulator *Thlaspi caerulescens*. Plant Physiology 142, 1127-1147

Van de Peer Y, De Wachter R (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Applic Biosc 10, 569-570

Vekemans X, Lefebvre C (1997) On the evolution of heavy-metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. Journal of Evolutionary Biology, 10:175-191

Xing JP, Jiang RF, Ueno D, Ma JF, Schat H, McGrath SP, Zhao FJ (2008) variation in root-to-shoot translocation of cadmium and zinc among different accessions of the hyperaccumulators *Thlaspi caerulescens* and *Thlaspi praecox*. New Phytologist 178, 315-325

Zha HG, Jiang RF, Zhao FJ, Vooijs R, Schat H, Barker JHA, McGrath SP (2004) Co-segregation analysis of cadmium and zinc accumulation in *Thlaspi caerulescens* interecotypic crosses. New Phytologist 163, 299-312

Zhao FJ, Hamon RE, Lombi E, McLaughlin MJ, McGraph SP (2002) Characteristics of cadmium uptake in two contrasting accessions of the hyperaccumulator *Thlaspi caerulescens*. J Exp Bot 53: 535-543

Chapter 5

General discussion

The hyperaccumulating species J. & C. Presl (Brassicaceae) has been chosen as one of the *Thlaspi caerulescens* models to study heavy metal tolerance and hyperaccumulation (Assunção *et al*., 2003a). It can hyperaccumulate cadmium (Cd), zinc (Zn) and nickel (Ni) and it is a close relative to *Arabidopsis thaliana* (L.) Heynh, with on average 88.5% DNA identity in coding regions (Rigola et al., 2006) and 87% DNA identity in the intergenic transcribed spacer regions (Peer *et al*., 2003). *T. caerulescens* is one of the few known Cd hyperaccumulating species (Baker *et al.*, 2000), but this trait is not found in all accessions of this species. For instance, the accession La Calamine (LC) is very tolerant to Cd but does not hyperaccumulate this element. This is in contrast to Ganges (GA) originating from southern France, which is both Cd tolerant and Cd hyperaccumulating (Reeves *et al.*, 2001; Assunção *et al*., 2003b). This pronounced intra-specific variation in the degrees and the metal-selectivity patterns of heavy metal hyperaccumulation offer excellent possibilities to identify underlying mechanisms at the molecular level through intra-specific comparisons and co-segregation analysis of crosses between accessions with contrasting phenotypes for any hyperaccumulation-associated traits.

The genetic variation among and within naturally occurring accessions of *Thlaspi caerulescens*, collected from different geographical regions (Assunção *et al*., 2003b) provides a source of genetic variation that can be used to study the function of genes. The accessions LC and GA exhibited natural genetic variation for all the traits described in this thesis. This variation was analysed using a F_2 population and its derived F_3 population specifically developed for this thesis and genotyped with

amplified fragment length polymorphism (AFLP) and other PCR based markers.

The main aim of the work presented in this thesis is to increase the knowledge of the genetic and physiological mechanisms underlying Cd hyperaccumulation in *Thlaspi caerulescens*. For that, we have analysed the intra-specific variation in the capacity to hyperaccumulate Cd but also Zn and Ni.

A continuous distribution of Zn and Cd accumulation was found for segregating populations derived from *T. caerulescens* intraspecific crosses (Assunção *et al.,* 2003b; Zha *et al.,* 2004, Assunção *et al.,* 2006). Such quantitative genetic variation has been exploited to detect and locate the loci contributing to the accumulation of Zn (Assunção *et al.,* 2006; Deniau *et al.,* 2006; Chapter 2 & 3), Cd (Deniau *et al.,* 2006; Chapter 2 & 3) and Ni (Chapter 3) using a so-called quantitative trait loci (QTL) analysis (Alonso-Blanco & Koornneef, 2000). Assunção *et al.* (2006) provided the first genetic map of *T. caerulescens*. No significant QTL were found to explain the observed variation for Zn accumulation in shoots in the segregating population, although the phenotypic frequency distribution suggests this trait to be heritable and polygenic. This might be due to the small size of the authors' mapping population. However, for Zn accumulation in the roots two QTL were found with, surprisingly, the trait-enhancing alleles coming from different parents. In addition, Chapter 2 focused on mapping the QTL for Cd and Zn accumulation in another segregating *Thlaspi caerulescens* population (LC x GA). The frequency distributions of the F_2 population for Cd or Zn concentration in shoot and root were not bimodal suggesting that Cd

and Zn accumulation inherited quantitatively. Similar findings were described in two other studies (Assunção *et al.,* 2003b; Zha *et al.,* 2004). The Zn and Cd accumulation in shoot and in root were significantly correlated, which was also found for the shoot concentrations in the study of Zha *et al.* (2004). In both cases the correlations were far from strict however, which could result from the contribution of mechanisms with different affinity patterns for both metals such as suggested by Zha *et al.* (2004). The identification of QTL, in Chapter 2, for Zn accumulation in shoot and in root with trait-enhancing alleles originating from both parents and the significant transgression found for Zn accumulation is in agreement with the hypothesis of two systems for Zn accumulation as suggested by Zha *et al.* (2004). This does not seem to be the case for Cd accumulation where the trait-enhancing alleles originate from the high Cd-accumulating accession. This could be due to the absence of the high affinity Cd uptake system in the low Cd-accumulating accession, such as postulated by Zha *et al.* (2004). In Chapter 2, the QTL analysis confirmed that all the studied traits, except for Cd concentration in shoot, are controlled by more than one gene. Interestingly, there was one case of co-location of QTL for Zn and Cd, with both trait-enhancing alleles originating from the high Cd-accumulating accession. These QTL explained most of the trait variances, both for Zn and Cd. This could be the result of a Zn/Cd transporter with relatively high affinity for Cd, such as proposed to be present in the high Cd-accumulating accession (Zha *et al.,* 2004). Such a transporter might explain the significant correlation between Zn and Cd accumulation. Taken together the results of Assunção *et al.* (2006) and

Chapter 2, it is evident that there are at least four loci determining the inter-accession variation in Zn accumulation in root. Dependent on the origin of the parents, either of these loci may or may not contribute to the segregation of the trait in inter-accession crosses.

In addition, the best way to further unravel the identified QTL is to first identify the corresponding regions of the Arabidopsis genome and use this information to either fine-map the region further or search directly for possible candidate genes based on a presumed function in metal homeostasis (Deniau *et al.,* 2006). Therefore, Chapter 3 focuses not only on mapping QTL for Zn, Cd and Ni hyperaccumulation in the same intra-specific cross, but also intends to present the first comparative mapping of *Thlaspi caerulescens* and the Brassicaceae framework. The number of orthologous markers mapped on the LC x GA genetic map enabled us to make further assumptions on the whole genome synteny and to establish chromosomal rearrangements as was done for another Zn and Cd hyperaccumulator *Arabidopsis halleri* (Willems *et al*., 2007) and the related *A. lyrata* (Kuittinen *et al.*, 2004; Koch and Kiefer, 2005; Yogeeswaran *et al.*, 2005). The *T. caerulescens* comparative map revealed that the Thlaspi species shared conserved segments that can be related to the ancestor karyotype (AK) (Chapter 3). The transition from eight chromosomes of the AK map to seven chromosomes in the *Thlaspi caerulescens* map can be explained by nine main chromosomal rearrangements (four translocations, two fusions and three inversions). This comparative genomic study will facilitate the further unravelling of the identified QTL. The QTL analysis did not confirm that all the studied traits, except for CdR and ZnR(Cd), are controlled by more than one

gene, as expected from the phenotypic segregation patterns. This suggests that traits concerned are governed by genes with little effect, so minor QTL still remain to be detected. We confirmed the detection of a major QTL for CdS on LG3. This QTL is located at the same position as the one detected in the F_2 population (Chapter 2). In the nearby region on LG3, a large cluster a QTL is found, grouping three QTL for Zn accumulation in root (ZnR, ZnR(Cd) and ZnR(Ni)), Cd accumulation in root QTL (CdR-1) and also the NiS QTL.

Remarkably, none of the genes involved in metal homeostasis and mapped on the F_3 populations are related to the QTL detected in this experiment. This indicates that heavy metal hyperaccumulation in *Thlaspi caerulescens* is governed by other genes or regulators. Several studies have revealed comparable expression levels in different hyperaccumulator accessions of genes that are strongly differentially expressed between hyperaccumulator and non-hyperaccumulator species (Assunção *et al.*, 1999; Xing *et al*., 2008), and genes that have been shown to be responsible for Cd and Zn tolerance QTL in a *A. halleri* x *A. lyrata* (hyperaccumulator x non-hyperaccumulator) cross (Willems *et al.*, 2007; Courbot *et al.*, 2007) are not necessarily differentially expressed between hyperaccumulator accessions with different accumulation and tolerance characteristics (Assunção *et al.*, 1999; Xing *et al*., 2008).

The present genetic maps and the accompanying mapping populations open new avenues for the identification of genes involved in the intra-specific variation of Zn, Cd and Ni hyperaccumulation in *T. caerulescens*. The QTL information offers new targets for investigating the molecular regulation of these traits. Further fine mapping will enable

a greater precision of QTL location and a more detailed synteny with the Brassicaceae genome framework.

Indentifying the genes responsible for the heavy metal hyperaccumulation traits and the footprints of selection acting on them are central to our understanding of the evolution of the trait. We observed a considerable variation in the degree of heavy metal hyperaccumulation between accessions when grown under controlled conditions in shoot and in root, respectively, for the three metals (Cd, Zn and Ni). These results are clearly in line with the previous studies from Assunção *et al.* (2003b) and Peer *et al,* (2003, 2006).

Contrasting with the strong differences in quantitative traits (Cd, Zn and Ni hyperaccumulation in shoot and in root, respectively), we found only little molecular variation among accessions, both for non-metal-related and metal-related markers. Overall, the non-synonymous and synonymous distances among Thlaspi accessions are not significantly different. The variation in non-coding regions was generally much higher than the variation in the coding regions. There was no significant difference between the non-synonymous and the synonymous distances with the corresponding sequences from *Arabidopsis thaliana*, except for NRAMP3, NAS4, MRP4 and ZNT1, where the synonymous distances significantly exceeded the non-synonymous ones. This is all in agreement with predominantly neutral substitution, or purifying selection in case of NRAMP3, NAS4, MRP4 and ZNT1.

The angiosperm family Brassicaceae (the mustard family) contains several important research models and agricultural crops, the foremost being the model species *Arabidopsis thaliana* (Arabidopsis) and the

Brassica crops. In addition, several related species are the focus of active research communities, including *A. lyrata, A. halleri*, *Capsella rubella*, and other genera such as Boechera, Lepidium, Thellungiella (also known as Eutrema) and Thlaspi. The recent advances in understanding the phylogenetics, polyploidization and comparative genomics in the family Brassicaceae paves the way to a unified comparative genomic framework (Schranz *et al.* 2006). This type of initiative is particularly timely because genome-sequencing projects are currently underway for *A. lyrata*, *C. rubella*, *T. halophila* and *B. rapa*; and since additional genetic mapping, cytogenetic and genomic data accumulate. It is to expected that the identification of candidate genes for any traits will be strongly facilitated by further comparative mapping and synteny analysis.

In conclusion, knowledge of the genes controlling specific steps of the metal homeostasis network is still rudimentary, but increasing rapidly. Increased knowledge on metal homeostasis is an urgent step needed right now to be able to better understand plant mineral acquisition, transport and storage. This will be facilitated by advances in high-throughput profiling of the transcriptome, proteome, metabolome and ionome (Salt, 2004). Such knowledge is expected to efficiently improve crop yield, crop nutritional value and food safety, three items which are of major global concern (Ghandilyan *et al.,* 2006). A multidisciplinary research effort that integrates the work of molecular and plant biologists, soil chemists and microbiologists is essential for better understanding of metal hyperaccumulation in plants.

References

Alonso-Blanco C, Koornneef M (2000) Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. Trends in Plant Science 5, 22–29

Assunção AGL, Schat H, Aarts MGM (2003a) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. New Phytologist 159, 411-419

Assunção AGL, Ten Bookum WM, Nelissen HJM, Vooijs R, Schat H, Ernst WHO (2003b) Differential metal-specific tolerance and accumulation patterns among *Thlaspi caerulescens* populations originating from different soil types. New Phytologist 159, 411-419

Assunção AGL, Pieper B, Vromans J, Lindhout P, Aarts MGM, Schat H (2006) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. New Phytologist 170, 21-32

Baker AJM, McGrath SP, Reeves DR, Smith JAC (2000) Metal hyperaccumulators plants: a review of the ecology and physiology of a biological resource for phytoremediation of metal-polluted soils. In: Terry N, Banuelos G, editors. Phytoremediation of contaminated soil and water. Boca Raton, FL: CRC Press LLC: pp 85-107

Boyd RS (2007) The defense hypothesis of elemental hyperaccumulation: status, challenges and new directions. Plant and Soil 293, 153-176

Courbot M, Willems G, Motte P, Arvidsson S, Roosens N, Saumitou-Laprade P, Verbruggen N (2007) A major quantitative trait locus

for cadmium tolerance in *Arabidopsis halleri* colocalizes with HMA4, a gene encoding a heavy metal ATPase. Plant Physiology 144, 1052-1065

Deniau AX, Pieper B, Ten Bookum WM, Lindhout P, Aarts MGM, Schat H (2006) QTL analysis of cadmium and zinc accumulation in the heavy metal hyperaccumulator *Thlaspi caerulescens*. Theoretical and Applied Genetics 113, 907–920

Ghandilyan A, Vreugdenhil D, Aarts MGM (2006) Progress in the genetic understanding of plant iron and zinc nutrition. Physiologia Plantarum 126, 407-417

Guimaraes MD, Gustin JL, Salt DE (2009) Reciprocal grafting separates the roles of the root and shoot in zinc hyperaccumulation in *Thlaspi caerulescens*. New Phytologist 184, 323-329

Hanikenne M, Talke IN, Haydon MJ, Lanz C, Nolte A, Motte P, Kroymann J, Weigel D, Kramer U (2008) Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. Nature 453, 391-U44.

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M, Mitchell-Olds T, Langley C, Savolainen O (2004) Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *Arabidopsis thaliana*. Genetics 168: 1575– 1584

Koch MA and Kiefer M (2005) Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of the three diploid species – *Capsella rubella*, *Arabidopsis lyrata* subsp *Petrea*, and *A. thaliana*. Am. J. Bot. 92: 761-767

Peer WA, Mamaoudian M, Lahner B, Reeves RD, Murphy AS, Salt DE (2003) Identifying model metal hyperaccumulating plants: germplasm analysis of 20 Brassicaceae accessions from a wide geographical area. New Phytologist 159, 421-430

Peer WA, Mahmoudian M, Freeman JL, Lahner B, Richards EL, Reeves RD, Murphy AS, Salt DE (2006) Assessment of plants from the Brassicaceae family as genetic models for the study of nickel and zinc hyperaccumulation. New Phytologist 172(2), 248-260

Richau KH, Kozhevnikova AD, Seregin IV, Vooijs R, Koevoets PLM, Smith JAC, Ivanov VB, Schat H (2009) Chelation by histidine inhibits the vacuolar sequestration of nickel in roots of the hyperaccumulator *Thlaspi caerulescens*. New Phytologist 183, 106-116

Rigola, D, Fiers M, Vurro E, Aarts MGM (2006) The heavy metal hyperaccumulator *Thlaspi caerulescens* expresses many species-specific genes as identified by comparative EST analysis. New Phytologist 170, 753-766

Salt DE (2004) Update on plant ionomics. Plant Physiology 136, 2451-2456

Schranz E, Lysak MA, Mitchell-Olds T (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer. Trends in Plant Science 11, 535-542

Verbruggen N, Hermans C, Schat H (2009) Molecular mechanisms of metal hyperaccumulation in plants. New Phytologist 182, 781-781

Willems G, Dräger DB, Courbot M, Godé C, Verbrugen N, Saumitou-Laprade P (2007) The genetic basis of zinc tolerance in

the metallophyte *Arabidopsis halleri* ssp. *halleri* (Brassicaceae): an analysis of quantitative trait loci. Genetics 176, 659 – 674

Xing JP, Jiang RF, Ueno D, Ma JF, Schat H, McGrath SP, Zhao FJ (2008) Variation in root-to-shoot translocation of cadmium and zinc among different accessions of the hyperaccumulators *Thlaspi caerulescens* and *Thlaspi praecox*. New Phytologist 178, 315-325

Yogeeswaran K, Frary KA, York TL, Amenta A, Lesser AH, Nasrallah JB, Tanksley SD, Nasrallah ME (2005) Comparative genome analyses of Arabidopsis spp.: Inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. Genome research 15, 505-515

Zha HG, Jiang RF, Zhao FJ, Vooijs R, Schat H, Barker JHA, McGrath SP (2004) Co-segregation analysis of cadmium and zinc accumulation in *Thlaspi caerulescens* interecotypic crosses. New Phytologist 163, 299-312

Summary

Thlaspi caerulescens (Tc; 2n=14) is a natural zinc (Zn), cadmium (Cd) and nickel (Ni) hyperaccumulator species belonging to the Brassicaceae family. It shares 88% DNA identity in the coding regions with *Arabidopsis thaliana* (At) (Rigola et al., 2006). It occurs both on metalliferous and non-metalliferous soils. Although the physiology of heavy metal (hyper)accumulation has been intensively studied, the molecular genetics are still largely unexplored. In Chapter 2, we address this topic by constructing a genetic map based on AFLP® markers and Expressed Sequence Tags (ESTs). To establish a genetic map, an F2 population of 129 individuals was generated from a cross between a plant from a Pb/Cd/Zn-contaminated site near La Calamine (LC), Belgium, and a plant from a comparable site near Ganges (GA), France. These two accessions show different degrees of Zn, Ni and particularly Cd accumulation. We analyzed 205 AFLP® markers (of which 4 co-dominant) and 12 co-dominant EST sequences-based markers and mapped them to seven linkage groups (LGs), presumably corresponding to the seven chromosomes of *T. caerulescens*. The total length of the genetic map is 516 cM with an average density of one marker every 2.5 cM. This map was used for Quantitative Trait Locus (QTL) mapping in the F2. For both metal concentration in shoot and in root we mapped two QTLs for Cd and Zn each. These QTLs explain 22 to 42% of the variance of the traits measured. We found only one common locus (LG3) for Cd and Zn (concentration in root). For all QTLs GA increased the trait value except one for Zn accumulation in shoot (LG1).

In Chapter 3, we constructed another genetic map based on AFLP® markers and ESTs and we compared the *Thlaspi caerulescens* map with the unified comparative genomic framework of the Brassicaceae. To establish a genetic map, an F3 population of 145 individuals was generated from the F2 of Chapter 2. We analyzed 170 AFLP markers and 44 co-dominant EST sequences-based markers and mapped them to seven linkage groups (LGs). The total length of the genetic map was 521 cM with an average density of one marker every 2.45 cM. The transition from eight chromosomes of the AK map to seven chromosomes in the *Thlaspi caerulescens* map can be explained by nine main chromosomal rearrangements (four translocations, two fusions and three inversions). For all traits except for Zn concentration in the shoot of plants under Cd exposure at least one QTL was detected. The 11 QTLs with significant effects were distributed over four LGs. Individual QTLs accounted for 8.5-40% of the phenotypic variation in this population.

Identifying the genes involved in heavy metal hyperaccumulation traits and evidencing the footprints of the natural selection acting on them are essential to understand the evolution of the trait. In Chapter 4, we observed a considerable variation in the degree of heavy metal hyperaccumulation in shoot and in root between *Thlaspi caerulescens* accessions for three metals (Cd, Zn and Ni). Contrasting with the strong differences in quantitative traits (Cd, Zn and Ni hyperaccumulation in shoot and in root, respectively), we barely found differentiation for metal-related and non-metal-related molecular markers. Phylogeny analysis identified two major geographic groups, i.e. the accessions from northwestern Europe which were all strongly related to each other and

the accessions from southern France which were less related among each other. Two isolated serpentine accessions from northern Italy and northern Spain were more distant from these groups as well as from each other. Genetic distances between accessions varied between markers but were overall too low to infer any conclusions on selection. Non-synonymous and synonymous distances to Arabidopsis varied between markers but were not significantly different from each other, except for *ZNT1*, *MRP4*, *NRAMP3* and *NAS4*, which showed an excess of synonymous substitutions, suggesting predominantly neutral genetic change and purifying selection on the latter markers, respectively. The patterns of genetic variation among accessions were not correlated with their metal accumulation behavior, nor with the degree and kind of soil metal enrichment at the site of origin. Adaptation to metaliferous soils appears to have occurred repeatedly at a local scale.

Samenvatting

De Zinkboerenkers, *Thlaspi caerulescens* (Tc; 2n = 14), is van nature een zink (Zn), cadmium (Cd) of nickel (Ni) hyperaccumulerende plantensoort, behorend tot de familie der Kruisbloemigen (Brassicaceae). De soort komt voor op bodems met een sterk verhoogd gehalte aan zware metalen, maar ook op 'normale' bodems. De basenvolgorde van het coderende DNA van deze plant is voor 88% identiek met dat van *Arabidopsis thaliana* (Rigola et al., 2006). Hoewel de fysiologie van hyperaccumulatie van zware metalen al decennia intensief bestudeerd wordt, is de (moleculaire) genetica van dit fenomeen nog grotendeels onontgonnen. In hoofdstuk 2 besteedden we aandacht aan dit onderwerp middels de constructie van een genetische kaart, gebaseerd op AFLP® merkers en "Expressed Sequence Tags" (ESTs). Om deze kaart te kunnen construeren werd een F_2 populatie van 129 individuen gegenereerd uit een kruising tussen een plant afkomstig van een sterk door Pb, Zn en Cd vervuilde plek bij La Calamine (LC), België, en één van een vergelijkbare groeiplaats in de buurt van het dorp Ganges (GA), in Zuid-Frankrijk. Deze twee locale populaties ("accessies") verschillen in de mate van accumulatie van Zn, Ni en, vooral, Cd. We hebben de recombinatie van 205 AFLP® merkers (waarvan 4 co-dominant) en 12 co-dominante, op ESTs gebaseerde merkers geanalyseerd. Deze merkers verdeelden zich over zeven "linkage groepen" (LGs), die vermoedelijk met de zeven chromosomen van *T. caerulescens* corresponderen. De totale lengte van alle LGs bij elkaar was 516 cM, met een gemiddelde merkerdichtheid van één per 2,5 cM. Deze kaart werd gebruikt voor het

karteren van “Quantitative Trait Loci” (QTLs), gebaseerd op de fenotypische variatie binnen de F_2 populatie. Voor de eigenschappen metaalaccumulatie in wortel en metallaccumulatie in spruit werden elk twee QTLs gevonden, zowel in het geval van Zn, als in het geval van Cd. Deze QTLs verklaarden 22 tot 42% van de totale fenotypische variatie van deze eigenschappen. We vonden slechts één gemeenschappelijke QTL voor Zn en Cd, namelijk voor accumulatie in de wortel. Van alle QTLs was het accumulatieverhogende allel afkomstig van de GA ouder, met uitzondering van één QTL voor Zn accumulatie in de spruit (LG1).

In hoofdstuk 3 construeerden we nog een genetische kaart, gebaseerd op AFLP® merkers en ESTs, en we vergeleken deze kaart met het “unified comparative genomic framework” van de Kruisbloemigen. Om deze kaart te maken werd een F_3 populatie van 145 individuen gegenereerd uit de F_2 populatie van hoofdstuk 2. We analyseerden 170 AFLP® merkers en 44 co-dominante, op ESTs gebaseerde merkers. Ook nu groepeerden deze merkers zich tot zeven “linkage groepen”. De totale lengte van de kaart was 521 cM met een gemiddelde dichtheid van één merker per 2,45 cM. De transitie van de acht chromosomen van de voorouderlijke AK (“ancestral karyotype”) kaart naar de zeven chromosomen van de *Thlaspi caerulescens* kaart zijn te verklaren door negen grote chromosomale reorganisaties, namnelijk vier translocaties, twee fusies en drie inversies. Voor elk van de eigenschappen Zn, Cd of Ni accumulatie in wortel of in spruit werd tenminste één QTL gedetecteerd. De in totaal 11 significante QTLs waren over vier LGs verspreid. Individuele QTLs verklaarden 8.5 – 40% van de totale fenotypische variatie binnen de F_3 populatie.

Het identificeren van de genen die betrokken zijn bij de hyperaccumulatie van zware metalen en het herkennen van de sporen van de natuurlijke selectie in de structuur van deze genen is van essentieel belang voor een goed begrip van de evolutie van het fenomeen. In hoofdstuk 4 observeerden we een aanzienlijke variatie tussen *T. caerulescens* accessies, wat betreft de mate van accumulatie van Zn, Cd of Ni in spruit en wortel. In schril contrast hiermee vonden we nauwelijks enige differentiatie van metaalgerelateerde en niet-metaalgerelateerde moleculaire markers. Met behulp van phylogenetische analyse konden twee duidelijk verschillende geografische groepen van accessies onderscheiden worden, namelijk die van Noordwest-Europa en die van Zuid-Frankrijk. De accessies van Noordwest- Europa waren onderling zeer verwant, veel meer dan die van Zuid-Frankrijk, ondanks de veel kleinere onderlinge geografische afstanden binnen de laatste groep. Twee geisoleerde accessies uit Noord-Italië en Noord-Spanje stonden tamelijk ver van deze beide groepen af, maar ook van elkaar. De genetische afstanden tussen accessies varieerden per merker, maar waren te gering om er eventuele sporen van natuurlijke selectie in te kunnen herkennen. De synonieme en niet-synonieme substitutiefrequenties tussen *T. caerulescens* en *A. thaliana* waren doorgaans niet significant verschillend van elkaar, behalve voor *ZNT1*, *MRP4*, *NRAMP3* en *NAS4*, die allen een relatieve overmaat aan synonieme substituties vertoonden. Deze resultaten duiden op selectief neutrale variatie, dan wel conserverende (‘purifying”) selectie, respectievelijk. De patronen van genetische variatie tussen de accessies waren niet gecorreleerd met die van de fenotypische variatie in

metallaccumulatie, noch met die van de mate of de aard van metaalverrijking van de bodem. Evolutionaire aanpassing aan sterk metaalhoudende bodems blijkt herhaalde malen, op locale schaal, opgetreden te zijn.

Acknowledgements

I cannot believe I did it!

When I was a teenager, I had a dream: to become a researcher. Well, I have tried the “royal way” in France but I was definitively not ready to go through the French theoretical education system. Then, after a difficult start (mainly for my parents!), I took a more practical road keeping my dream in mind. Later I took the opportunity to come study in The Netherlands and to get into the Dutch education system: “learning by doing”. This fitted me much better and stimulated me to reset my professional goals. Then, plant breeding was clearly the field I wanted to play a role in. The opportunity to do a first step in academic research came when I did a first internship in Paul Scott’s lab in the university of Iowa State. This experience really showed me that my dream could become reality. There, I met a French breeder (I forgot his name) who told me that a “detour” by academic research would help for my professional career in plant breeding. Freshly graduated for a HBO, my mind was set to do a MSc at Wageningen University so, in order to efficiently use the time I got to wait before the beginning of the MSc, I took contact with Maarteen Koornneff, and I got the opportunity to learn more in Mark Aarts’ lab. At the end of this voluntary internship, Mark was complaining that he could not find a PhD student for a project he got in collaboration with the famous Henk Schat. Diana Rigola, a post-doc in Mark’s lab, proposed I could do the job. I applied and I got the position. My dream was really starting to take shape. Couple of years later, I am

writing these words so it means I am becoming a researcher and my teenager's dream became true. Well, becoming a researcher … I am not so sure anymore. Between this day of August 2002 and today, the road was long. Sometimes (often?), it was paved of difficulties, frustrations, humiliations, sleepless nights. Is there not a saying: "The goal is not important but the way you take to achieve it"? Few times, I thought I met my limits, but thanks to a bunch of people, I stood up and got back on the road. Today, I am still amazed how my brain has forgotten all these dark moments and has kept in memory only the joy, fun, spirit that I enjoyed with the load of people I am about thanks.

My first thoughts go to the department of Ecology and Physiology of Plants at the Vrije Universiteit. Henk, you guided me all over this part of life. I am deeply thankful. Also, I am honored I met the stereotype of the researcher from my teenage hood! Veerle and Kerstin, I had so much fun working with you. We had great moments and it would take a book to file them all. Ana, Petra, (even) Emma, Jos, Henk, Riet, Thijs, Wilma thanks a lot for making this group a family. Our Italian visiting scientists, nice meeting you Barbara (and Andrea) and Miluscia.

Also, I would like to sincerely thank the famous student bar: "The Stelling". I have spent some of my greatest parties there with wonderful people.

In Wageningen, Mark is the first one I think of. I became your first student as you were just started your research group. You are a talented researcher but not such a good group leader! Thanks to Judith (great collaboration), Sangita, Artak, Diana.

At Plant Breeding, WUR, special dedicace to Benoit. I loved working with Petra, Fien, Pim and Sjaak.

To Pierre and Glenda. It was really nice to share our results. Merci!

Louis, you have been my only student during this Phd. You did a real good job on Chapter 4. It seems my experience did not convince you not to follow the same steps. You are determined to have a go. Good luck in Montpellier!

The French mafia in the Netherlands: my bodies Florent and Guillaume, you were always teasing me but, look at that: I did it! You know how to call me from now on!

A good word to Dr.Jair, Dr.Csaba and almost-Dr.Martijn from Rijk Zwaan. You kept me motivated on the last straight line.

To Ronald Koes, you make it to the acknowledgments. Technically, I could not promote without you. Thanks for your support.

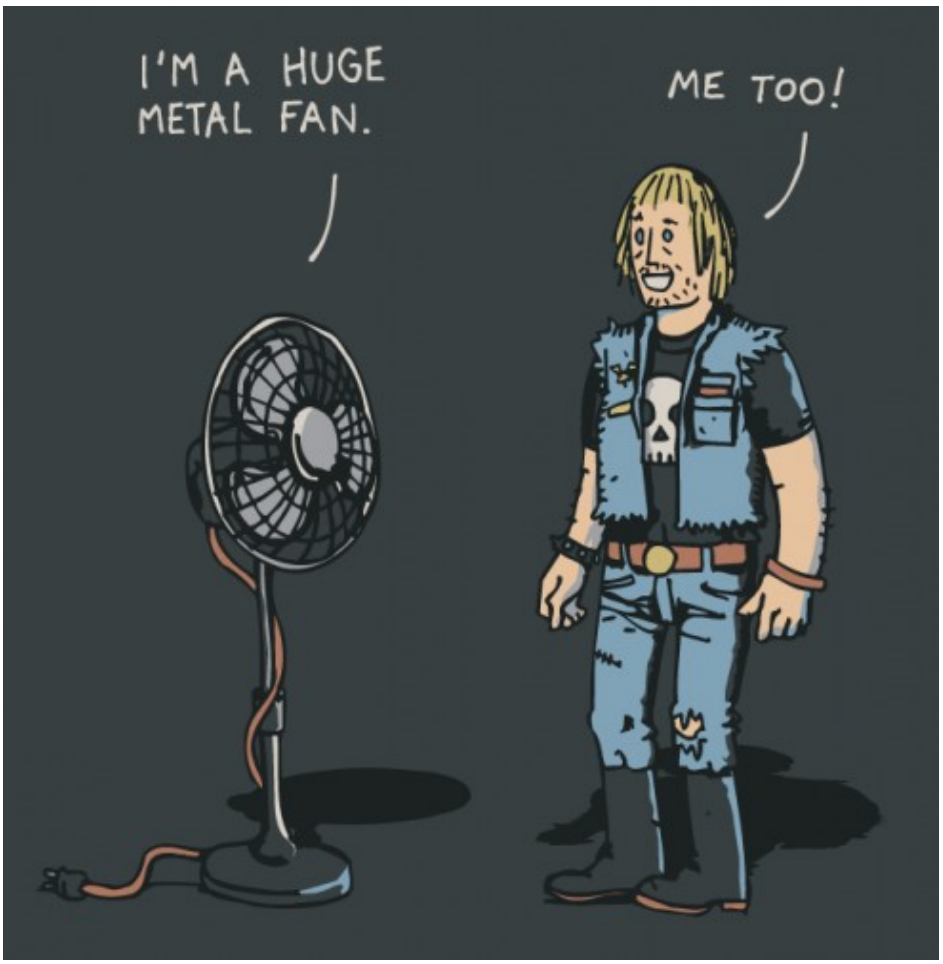
My family. I would always be grateful to my parents. You have always supported me in any circumstances. My sisters, Claire and Marion too. The only way I can thank you is to try to do the same for my own children (what a challenge!).

Last but not definitively not the least, Sandra, my beloved wife. I know it has not been easy but your daily support and love made it possible. I love you.

Bibliography

Deniau AX, Pieper B, Ten Bookum WM, Lindhout P, Aarts MGM, Schat H (2006) QTL analysis of cadmium and zinc accumulation in the heavy metal hyperaccumulator *Thlaspi caerulescens*. Theoretical and Applied Genetics 113, 907–920

Deniau AX, Schat H, Aarts MGM (2007) Genetics and genomics of the heavy metal hyperaccumulator model species *Thlaspi caerulescens*. Genes, Genomes and Genomics 1, 81-88


I'M A HUGE METAL FAN.
ME TOO!